2007

Approximate dynamic programming and aerial refueling

Panos, Dennis C.

http://hdl.handle.net/10945/2990

APPROXIMATE DYNAMIC PROGRAMMING AND AERIAL REFUELING

Dennis C. Panos

A THESIS
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF MASTER OF SCIENCE IN ENGINEERING

RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF OPERATIONS RESEARCH AND FINANCIAL ENGINEERING

JUNE 2007

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
XX-06-2007	Master's Thesis	JAN-JUN 2007

4. TITLE AND SUBTITLE
Approximate Dynamic Programming And Aerial Refueling

5a. CONTRACT NUMBER
N00244-99-G-0019

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Dennis C. Panos

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Princeton University

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Naval Postgraduate School
Monterey, Ca 93943

10. SPONSOR/MONITOR'S ACRONYM(S)
NPS

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
1. DISTRIBUTI9ON STATEMENT A. Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Aerial refueling is an integral part of the United States military's ability to strike targets around the world with an overwhelming and continuous projection of force. However, with an aging fleet of refueling tankers and an indefinite replacement schedule the optimization of tanker usage is vital to national security. Optimizing tanker and receiver refueling operations is a complicated endeavor as it can involve over a thousand of missions during a 24 hour period, as in Operation Iraqi Freedom and Operation Enduring Freedom. Therefore, a planning model which increases receiver mission capability, while reducing demands on tankers, can be used by the military to extend the capabilities of the current tanker fleet.

Aerial refueling optimization software, created in CASTLE Laboratory, solves the aerial refueling problem through a multi-period approximation dynamic programming approach. The multi-period approach is built around sequential linear programs, which incorporate value functions, to find the optimal refueling tracks for receivers and tankers. The use of value functions allows for a solution which optimizes over the entire horizon of the planning period. This approach varies greatly from the myopic optimization currently in use by the Air Force and produces superior results.

The aerial refueling model produces fast, consistent, robust results which require fewer tankers than current planning methods. The results are flexible enough to incorporate stochastic inputs, such as: varying refueling times and receiver mission loads, while still meeting all receiver refueling requirements. The model's ability to handle real world uncertainties while optimizing better than current methods provides a great leap forward in aerial refueling optimization.

The aerial refueling model, created in CASTLE Lab, can extend the capabilities of the current tanker fleet. Additionally, the robust nature of the aerial refueling model's solutions provides insight into the strength and flexibility of the approximate dynamic programming method.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					Sean Tibbitts, Educational Technician
a. REPORT	b. ABSTRACT	c. THIS PAGE	UU	141	19b. TELEPHONE NUMBER *(include area code)* (831) 656-2319 civins@nps.edu

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

ABSTRACT

Aerial refueling is an integral part of the United States military's ability to strike targets around the world with an overwhelming and continuous projection of force. However, with an aging fleet of refueling tankers and an indefinite replacement schedule the optimization of tanker usage is vital to national security. Optimizing tanker and receiver refueling operations is a complicated endeavor as it can involve over a thousand of missions during a 24 hour period, as in Operation Iraqi Freedom and Operation Enduring Freedom. Therefore, a planning model which increases receiver mission capability, while reducing demands on tankers, can be used by the military to extend the capabilities of the current tanker fleet.

Aerial refueling optimization software, created in CASTLE Laboratory, solves the aerial refueling problem through a multi-period approximation dynamic programming approach. The multi-period approach is built around sequential linear programs, which incorporate value functions, to find the optimal refueling tracks for receivers and tankers. The use of value functions allows for a solution which optimizes over the entire horizon of the planning period. This approach varies greatly from the myopic optimization currently in use by the Air Force and produces superior results.

The aerial refueling model produces fast, consistent, robust results which require fewer tankers than current planning methods. The results are flexible enough to incorporate stochastic inputs, such as: varying refueling times and receiver mission loads, while still meeting all receiver refueling requirements. The model's ability to handle real world uncertainties while optimizing better than current methods provides a great leap forward in aerial refueling optimization.

The aerial refueling model, created in CASTLE Lab, can extend the capabilities of the current tanker fleet. Additionally, the robust nature of the aerial refueling model's solutions provides insight into the strength and flexibility of the approximate dynamic programming method.

Contents

List of Figures

List of Tables

1 *Introduction*

A tenant of the doctrine guiding the modern United States military states that military forces need to respond around the world in a rapid manner with an overwhelming and continuous projection of force (7). Given the current geopolitical climate, the stated goals of a rapid response force which is both overwhelming in power and is able to operate over an extended time frame appear to be contradictory objectives. During the Cold War the United States was able to focus its assets on the former Soviet Union with forward deployed assets placed in Germany, Japan, South Korea, and other strategic locations which surrounded the Soviet Union. Therefore, through forward basing the United States' military was guaranteed the ability to respond rapidly and sustain a continued projection of force. However, since the fall of the Soviet Union and its satellites the political climate and requirements facing the United State miliary have become much less stable.

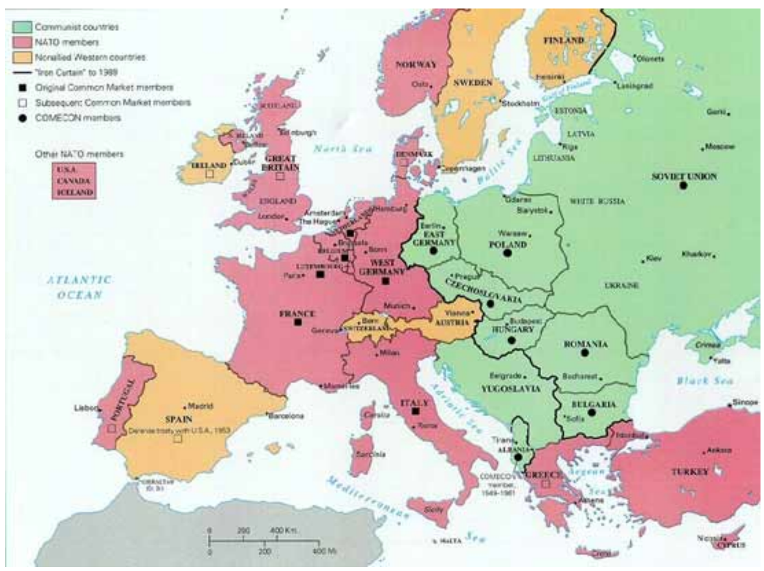

Figure 1: Map of the Political Climate of the Cold War

Due to the instability of the current political environment, the future requirements placed on the United States military cannot be guaranteed with any more accuracy than the fall of the Soviet Union was predicted. Additionally, while forward basing of United States troops on foreign soil was feasible during the Cold War, today other countries are far less accepting of having American troops stationed on their soil.

Lacking a definable future enemy and the ability to forward deploy troops around the globe, how does the United States expect to quickly respond to crises around the world with a mass of overwhelming and continued force?

Figure 2: Branches of American Military

The answer lies in the structure of the four branches of the American military. The modern Marine Corps is designed to respond rapidly and deploy short term ground assets around the world. The sustainment of the ground forces is the responsibility of the Army, which has the capability to follow the Marine Corps with a large force designed for continuous deployment. The shortcoming of the modern military is its ability to attack over the horizon with aerial assets due to the lack of forward basing.

The United States Navy has the ability to quickly traverse the oceans and operate in the littoral regions. The ability to work within close proximity to coastal nations allows the Navy to send ordinance deep into enemy territory. However, bombardment by Tomahawk missiles and projectiles is not the overwhelming force the United States military desires for over-the-horizon operations. It is through the joint efforts of the United States Air Force and Navy's aircraft inventory that the United States can gain both air superiority and the ability to send large masses of ordinance deep into enemy terrain.

Without forward basing, challenges exist such that the Air Force's aircraft inventory can be out of range of the belligerent nation and the Navy's aircraft also have limited ranges and cannot fly much further than the borders of large countries. Aerial refueling tankers with their extended range and fuel carrying capabilities provide a gas station in the sky and ensure longer ranges and time on station for other American aircraft. Through aerial refueling the Air Force and Navy are able to provide over-the-horizon power projection and air superiority which guarantees the American military's ability to rapidly respond around the world with an overwhelming and continuous projection of force.

1.1 *Aerial Refueling Background*

Mid-air refueling is both a technical challenge as well as a complex planning process. The highly orchestrated maneuvers required to refuel planes flying in excess of 300 knots per hour are multiplied as the Air Force inventory of mid-air refueling planes must refuel a variety of planes and helicopters flown by the Air Force, Navy and Marines. In addition to the technical challenges posed by refueling a myriad of different platforms, the planning of mid-air refueling in an incredibly complex process which always must weigh several different objectives. The military combat commander's desire to deliver ordinance on specific targets, at specified times, with an overwhelming mass of force, places great requirements on the air refueling assets. The overwhelming force requirement places large stresses on the aerial refueling fleet as missions often involve multiple aircraft, and the aircraft all require simultaneous refueling. The requirements are made even more acute due to bomber and attack/fighters planes ranges, which are often much shorter than the length of their missions. Additionally, hostile air space can limit the ability of aerial refueling tankers to escort attack planes to their targets. Therefore, in the modern era, the planning of aerial refueling is a major factor in determining mission success and the military's ability to operate efficiently.

Figure 3: KC-10 refueling the Joint Strike Fighter

1.2 *The Beginning of Aerial Refueling*

Mid-air refueling was not always such a highly integrated part of a military's battlefield success. During War World I an aircraft's effectiveness focused solely on the pilot's ability to shoot down the enemy and not a complex refueling scheme. Since no in-flight refueling protocol existed every plane in the air had limited range and time in the air. Surprisingly, this did not provide the impetus for the first attempts at aerial refueling. Rather, a vaudevillian act by a stunt man and a Naval Lieutenant years after the war, in 1921, was the first recorded "aerial refueling". In the first aerial refueling a stunt man walked out on the wing of a JN-4 plane and onto the wing of an adjacent JN-4 with a can of gas strapped to his back which he poured into the gas tank (5). Another early attempt, also in 1921, involved a Naval Lieutenant flying down the Potomac River and picking up a floating gas can with a grappling hook (19). While these attempts were very daring they did not provide insight into the problem of refueling while flying, unless of course the Navy started hiring circus performers or fisherman.

Two years later, in 1923, the first modern approach of a mid-air refueling using hoses passed between planes was successfully attempted by two Army Air Corps de Havilland DH-4Bs (9). While crude by modern standards, the passing of hoses between planes is effectively the same approach used over 80 years later. The early excitement generated by the Army's refueling example led to both an emerging commercial interest and a new breed of stunt men who became interested in aerial refueling. The Key brothers extended flight in 1935 provides an example of the length daredevils went to prove their machismo and the ability of planes to remain aloft semi permanently. While the brothers didn't walk on wings they used mid-air refueling to stay aloft for 27 straight days. During their flight, which remains a record to this day, they were resupplied through a primitive hose method 484 times, which clearly demonstrated the huge potential for mid-air refueling (9). The commercial sectors use of aerial refueling before World War II expanded through the interest of Shell Oil Company which owned the major producer of refueling hardware, Flight Refuel-

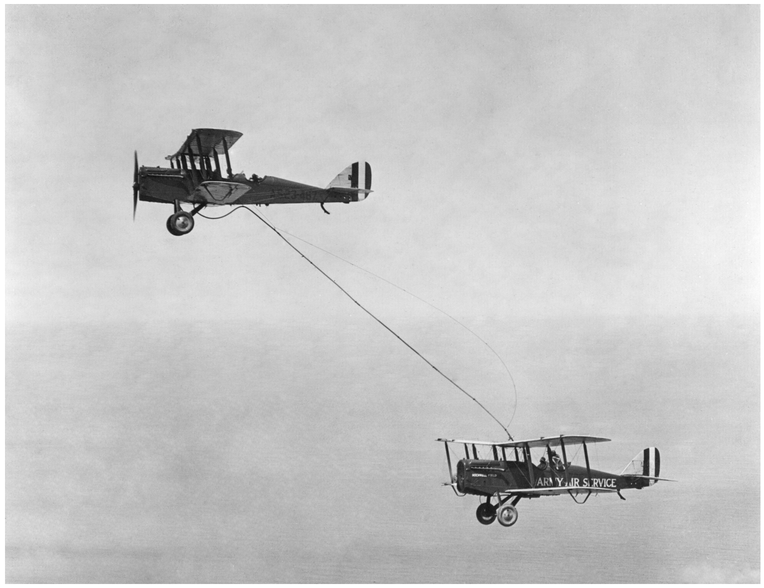

Figure 4: Aerial Refueling circa 1923

ing Limited (20). Shell Oil saw the sky as the limit for selling gasoline, and aerial refueling was used for transatlantic flights and mail routes.

Interestingly, the early demonstrations of the endurance enabled through in-flight refueling were not enough to see in-flight refueling enter World War II. The air battles fought in the Pacific would have benefitted through aerial refueling. Also aerial refueling would have enhanced the ability of the US miliary to attack German land targets; however, while the Army Air Corps and the US Navy continued research during World War II, they did not implement any of their aerial refueling knowledge.

An example of how World War II planners dismissed the idea of in-flight refueling was shown through their insistence that the military gain a foothold on Tinian Island in the Northern Marianas. The planners required Tinian so that they could construct an airfield which would allow the existing long range bomber in the American inventory, the B-2, to reach Japan and return unrefueled. It was not until the advent of the Cold War and the Nuclear Age that the strategic planning of the military ushered in the next chapter of aerial refueling.

1.3 *The Modernization of Aerial Refueling*

The atmosphere of fear and suspicion that surrounded the beginning of the Nuclear Age and Cold War brought forth great advancements in aerial refueling. Before the introduction of the Intercontinental Ballistic Missile the only way to deliver a nuclear payload on the Soviet Union was through Air Force and Naval bombers. With the extreme distances involved in reaching all points within the Soviet Union, aerial refueling was the only option for returning bombers after dropping their payloads. This lead to the Air Force demonstrating in 1949 that they could circumvent the world using aerial refueling (13). The mission, completed by a B-50A, involved 4 refuelings using a wire and hose system. While the mission was a success it still involved a highly specialized skill set, as it required a harpoon gun to fire linking wire between the planes, and the refueling was tedious and time consuming due to the limit on fuel flow through flexible hoses.

Figure 5: Boeing B-50A Superfortress.

Refining the method so that it was both easier and faster was a priority for the Air Force ,and they found a solution in the form of the American System, developed by Boeing (20),(9). The American System employed a semi-rigid, telescoping, swiveling refueling hose mounted to the fuselage of the refueling tanker, and the system also employed winged control surfaces for greater hose stability. With the American System the maneuvering required by the receiver plane during refueling was significantly decreased as greater control of the hose was afforded to the hose operator located on the tanker. Another improvement of the American System was the rate at which the fuel was transferred between the tanker and the receiver, which was much faster then the previous hoses systems. While there have been improvements to the American System, the foundation of system currently employed was introduced by Boeing in 1948. Since then the major changes to aerial refueling have focused upon tanker design and fleet size (2).

Figure 6: Lockheed C-5 Galaxy refueling by KC-135 with an Example of a Boom

In addition to improved refueling methods, the Cold War also necessitated a much larger fleet of tankers with increased capability due to the introduction of the Strategic Air Command (SAC). SAC was designed with the dual purpose of protecting the United States borders in cases of imminent attack from the Soviet Union and the rapid deployment of every asset capable of carrying a nuclear weapon into the Soviet Union. The greatest problem for SAC involved infiltrating the Soviet air space, since the introduction of the jet age made the propeller driven B-29 and B-52 bombers obsolete as Soviet fighters could easily catch these planes. The Air Force responded in 1954 by introduced the first long range jet bomber by retrofitting the B-52 with eight turbojet engines (9). The Air Force tested the capability of the retrofitted tankers during Operation Power Flite. The operation proved to be a success as it reduced the amount of time required to circumnavigate the earth to 45 hours, which was less than half time of the previously held record.

Operation Power Flite also highlighted a major deficiency of the jet powered tankers. When the tankers were flown outside their optimal speed and altitude they were highly inefficient. This deficiency was exacerbated by the fact that the air refueling planes at the time were turbo props and therefore required that the B-52s fly slow and low to refuel. Thus, the planes meant to extend the range of the jet bombers actually were also limiting the range of a fully refueled jet powered B-52. The next step for the Air Force was to find a suitable jet powered refueling plane so that the jet powered bombers could operate efficiently and reach their targets faster.

The competition to produce a jet powered tanker pitted Boeing against McDonnell Douglas and Lockheed Martin. In the competition Boeing took the early lead as the company possessed both a design and a working prototype (20). The Boeing design of the KC-135 Stratotanker was a working prototype which was based on the air frame of the Boeing 707. Given the urgency of the Cold War the Air Force adopted the KC-135. However, the KC-135 was adopted as an interim tanker, since even at its adoption the Air Force leaders had judged the other companies' designs to be superior to the Stratotanker.

After adopting the Stratotanker the mission planners were immediately faced with a tough refueling challenge. The lessons learned from Operation Power Flite showed the planners that for optimal deployment every B-52 produced would require a tanker in a one to one ratio. The rapid production of the B-52 in the mid 1950's necessitated the equal production of jet tankers so the KC-135 dropped its interim status and became the tanker of the United States Air Force. At the end of the production of the B-52 and KC-135 in the mid 1960s, 732 KC-135's had been produced and stationed around the United States. In spite of being judged the inferior design the KC-135 represented the introduction of the modern aerial refueling fleet for the US Air Force. The KC-135 has proven to be an incredibly durable airframe and continues its service in the US Air Force inventory today with avionics and engine retrofits. While other refueling platforms have been introduced, it was in the late 1950's that the modern equipment and methods of aerial refueling were finally introduced. However, it would take a change in a different type of technology for the modern aerial refueling mission to come into existence.

SAC depended heavily on the KC-135 for refueling long range jet bombers and fighters until the requirement of long range bombers changed drastically with the introduction of the ICBM. The reduction of the importance of the long range bomber curtailed the strategic need for jet tankers and their refueling capabilities. The mission of the aerial refueling fleet languished until the Vietnam War and a refocusing of the scope of the aerial refueling capabilities. Before the war the aerial refueling doctrine focused upon fueling bombers and fighters on their way to engagement and on their return from their engagement. In Vietnam, the mission of combat support was added as planes low on fuel during missions would refuel over the skies of Vietnam and resume their missions (20). This change was a shift in ideology from each receiver aircraft being paired with a specific tanker to the idea that each tanker could support a variety of planes and missions in a combat environment.

The Vietnam War also saw the first use of the hose and drogue system for refueling receivers. The hose and drogue system varies from the American System, also known as the boom system, in that there is a flexible hose with a cone attached which is dragged behind a tanker. With the hose and drogue system the receiver aircraft must fly their refueling point into the cone. Before the Vietnam War the hose and drogue system was implemented by the US Navy for its fighters and its helicopters and was used by the Navy's small refueling platform: the KA-3 tanker. As shown in Figure 7, the flexible hose can accommodate varying platforms such as helicopters while the fixed boom can not.

Since the Navy's planes were designed to accept the hose and drogue and not the boom system, the Air Force tankers could not refuel Naval assets. Additionally, the Air Force tankers where prohibited by SAC from refueling any non Air Force planes. However, ingenuity reigned the day and Air Force tankers frequently refueled Navy fighters (9). The tankers did so in indirect manner, as they could refuel KA-3 tankers with their boom system and the KA-3 would simultaneously or subsequently refuel Navy fighter/bombers with their hose and drogue system. Since the Vietnam War, as intra service cooperation has improved, the system of indirect fueling has been replaced by Air Force tankers being both boom and hose and drogue capable.

Figure 7: Example of Hose and Drogue

After the Vietnam War there have been exciting examples of how aerial refueling allows the prosecution of warfare and limited strikes on targets without forward basing. These examples laid the foundation for the creation of the modern mission capability. The first example of a long distance strike on a foreign target was performed during the British attack during at attack on the Falkland Islands in 1982. The British operation dubbed "Operation Black Buck" was a series of six long range bombing missions performed by the Royal Air Force Vulcan long range bomber(1). During the first mission two Vulcan aircraft were deployed from Wideawake airfield on the Ascension Islands more the 3,900 miles from their target at Port Stanley, Falkland Island. The Vulcan bomber developed in 1960, was designed to carry nuclear weapons within the confines of European soil and was therefore not suited for the long distance this mission required. With a quickly devised refueling strategy the Vulcan took off with a complement of eleven refueling aircraft. During the outbound flight the Vulcan was refueled five times, but more impressively there was tanker to tanker refueling which allowed the refueling procedure to cross the Atlantic. On the inbound flight the Vulcan only required one refueling which was all the tankers could provide as all the planes barely had enough fuel to return to the Ascension Islands. At the time of the attack the missions of "Operation Black Buck" were the longest combat mission flights in history and showed that, if necessary, in-flight refueling could allow aircraft to strike anywhere in the world. Figure 8 shows both the distances involved in refueling the Vulcan as well as the complexity of the refueling operations which included both tanker-receiver and tanker-tanker refueling.

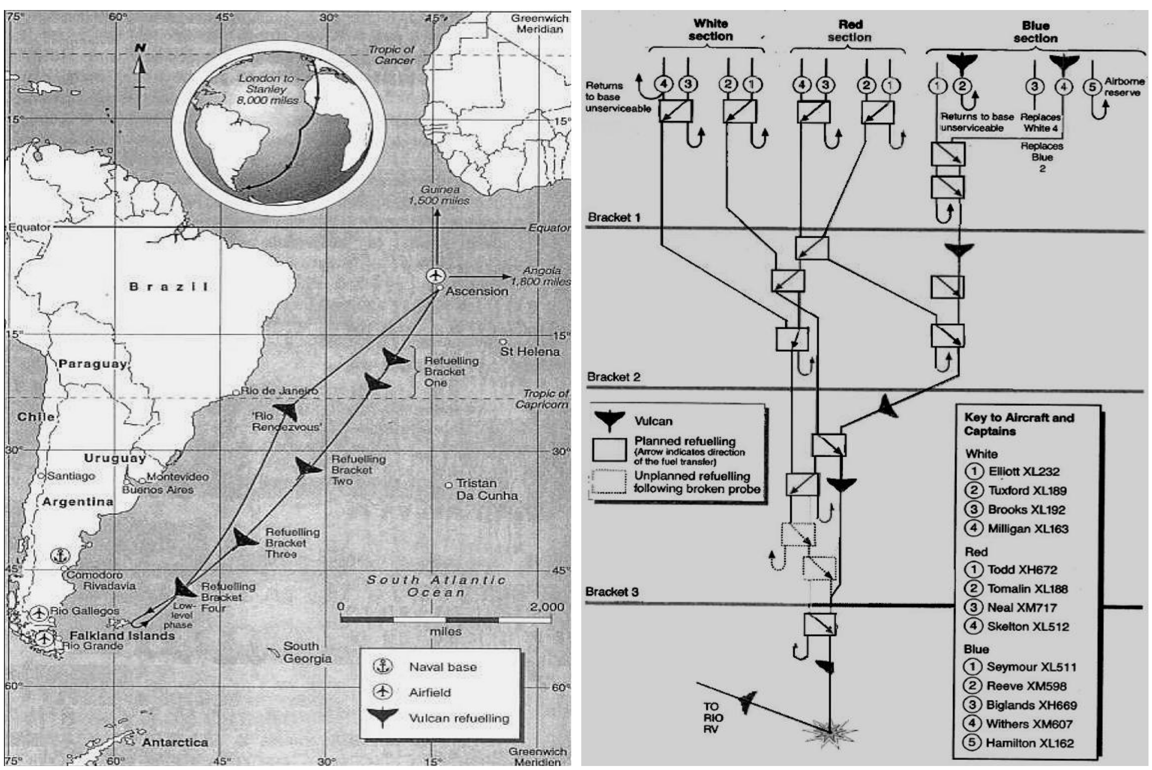

Figure 8: Operation Black Buck Refueling Schematics (22)

The United States also demonstrated its ability to prosecute long distance attacks using aerial refueling when it struck Libya after the acts of terrorism perpetrated by that state and its leader Muammar al-Qaddafi. While the British used eleven tankers to support one Vulcan bomber (1), the United States was able to limit that number through the use of the new KC-10 tanker.

The KC-10 Extender tanker was brought into service in 1981 and its capabilities far exceed that of the KC-135. The KC-10 has twice the fuel capacity of the KC-135, can employ both boom and hose and drogue systems, and can receive aerial refueling. The long distance strike against Libya (Operation El Dorado Canyon) was necessitated by the French refusal to grant overfly rights and thus direct routes against Libya were not an option from current US air bases (20). In a mission requiring much planning, the US took off from Mildenhall Air Force base in the United Kingdom with 24 F-111 fighters supported by 19 KC-10s which were subsequently supported by 10 KC-135s. The operation proved a success and showed that the United States could use aerial refueling to support rapid strikes on foreign targets with a mass of force, in addition to the missions previously defined.

1.4 *Modern Aerial Refueling and the Future*

The last 15 years have presented unique challenges to the aerial refueling community that could have never been anticipated by the first wing-walking refueler. The enormity of the missions flown in Operation Desert Storm placed challenges on the tanker fleet never before faced and highlighted the shortcomings of aerial refueling in a modern war. Additionally, while prosecuting targets in Afghanistan during Operation Enduring Freedom, aerial refueling faced the challenge of incredible mission distances and large mission loads.

1.4.1 *Desert Storm and Enduring Freedom*

Operation Desert Storm utilized both the combat operations and long distance support roles of aerial refueling. In 1990, when Iraq invaded Kuwait and massed on the Saudi Arabian border, there was a need for rapid deployment of troops and material, as well as a rapid response with military force against Iraq. While the internationally imposed deadlines for Iraqi withdrawal drew near, the United States created an “air bridge” to transport material and troops across the Atlantic and Pacific Oceans (12). These “air bridges” were actually C-5 and C-141 transport planes supported by over 100 tankers that transported required manpower and material from Europe and the United States to the Saudi Arabian airbases. The “air bridge” concept was successful because of the ability of tankers to refuel planes enroute without having to reroute loaded transport planes on longer routes that would have required the downtime of landing to refuel.

The United States also incorporated its improved concept of supporting long distance strikes during Operation Desert Storm. After the deadline for withdrawal passed the United States sent seven B-52 bombers loaded with cruise missiles from Barksdale Air Force Base in Louisiana (10). The seven planes refueled four times on the way to bombing targets in Baghdad which to that point was the longest strike in history.

The third and most integral part of the refueling mission in Operation Desert Storm focused on the combat refueling role played in and around the Iraqi airspace. The first conflict in Iraq involved the most tankers of any operation in history; which when combined with the number of sorties and the relatively small theater of operation constituted a major restructuring in how refueling was conducted (14). The close proximity of Saudi Arabian air bases where the tankers were forward based, along with the air superiority gained in the first weeks of the war, allowed tankers to work as an active refueling point for many receiver aircraft from both the Air Force and the Navy. In this role the tankers were able to get on station quickly, offload a maximum amount of fuel, and subsequently return to base and refuel themselves in a compressed time frame (14). This had not been the case in Vietnam when combat refueling was in its infancy or during the other long range escort refueling missions such as Libya. While theoretically the quick turn around time and the large amount of fuel that tankers could offload would be a boon to efficiency of missions and tanker usage, this was not the case.

Figure 9: The skies above Iraq - Operation Desert Storm

While the aerial refueling assets contributed mightily to the success of the air campaign, several studies by RAND and the GAO highlighted the shortcomings of the aerial refueling campaign. A GAO report states that "because of the finite amount of Saudi Arabian airspace and the large number of missions being supported each day, tanker refueling operations were frequently constrained by congestion" (15). Obviously that statement is of great concern as through improved efficiency comes the ability to prosecute a war more effectively. The questions posed were "why were there so many tankers in the air" and "were all the tankers required?". The GAO found that on average over 40 percent of the fuel a tanker took off with was unused by the end of the mission. They stated that the inefficiency of the operations limited additional combat missions since it appeared as though tankers were being assigned in the most conservative manner possible (15). The conservative approach of assigning tankers as needed to missions without regard for future needs or the current inventory of tankers in the air drew the ire of the RAND study which stated: "In the absence of automated planning tools, planners used planning factors to estimate the number of tankers in order to ensure mission success . . Better planning tools and training could conceivably result in great savings in required tanker sorties during major operations." (11). While a GAO study found that fuel returned to base decreased throughout the war due to better planning and utilization of assets in the sky, it was not due to official policy changes but rather operational planners learning on the job;

however, as the war finished this knowledge retired with the planners. While the war was a success and the capabilities enabled by aerial refueling played a major role it also highlighted shortcomings in the planning abilities of operational planners.

	Iraq(1991)	Kosovo(1999)	Afghanistan(2001-02)	Iraq(2003)
Aircraft	306	175	80	185
Sorties	16,865	5,215	15,468	6,193
Flight Hours	66,238	52,390	115,417	NA
Sorties/Hour	3.9	10.0	7.5	NA
Receiver Aircraft	51,696	23,095	50,585	28,899
Fuel off-loaded(lbs)	800.7M	253.8M	1,166M	376.4M
Av Fuel Sortie(lbs)	47.5K	48.7K	75.4K	60.8K

Table 1: Source: GAO analysis of Air Force Data

The latest test of American aerial refueling capabilities came during Operation Enduring Freedom. During Operation Enduring Freedom, the capability of air refueling assets to help prosecute a war over great distances was severely tested. The distance traveled to and from targets within Afghanistan rivaled those of the long distance strikes accomplished in the past; however, they were not single isolated strikes but rather continuous strikes across the country in support of a war. Given the landscape and political climate in southwest Asia the coalition assets had to fly from aircraft carriers distances of over 700 miles or from the British protectorate of Diego Garcia more than 3000 miles away. Additionally, with the inclusion of the B-2 bomber in the US arsenal, 30 hour missions covering half the globe were also used for covert operations (8). The complexity of missions which involved great distances, the continued need for planes attacking both fixed targets as well targets of opportunity, and close air support required better planning than ever before. During Operation Enduring Freedom the sortie rates were in line with the amount in Desert Storm shown in Table 1. However, in Operation Enduring Freedom each offload was nearly 40 percent larger than those in Desert Storm, and the sortie lengths were much longer and therefore receivers required multiple refuelings per sortie. The war in Afghanistan highlighted the reliance of modern warfare on aerial refueling and the current American capacity to meet that reliance.

1.4.2 *The Future*

The US tanker fleet is an aging fleet with major components made up with hold-over KC-135s from the early 1960s (4). In the past several years there have been studies researching the need for new tankers with better range, more fuel capacity, and the ability to refuel more than one receiver at a time (2). These studies have focused on the aging fleet and the requirements placed on the tanker fleet over the past 15 years. Adding the ability to refuel multiple aircraft simultaneously through multi point refueling stations is a way to get around the under-utilization of tankers from the first Gulf War. The possibility that a future belligerent nation will be a long distance from any forward base or the ocean highlights the need for both more tankers as well as more reliable tankers (3). The government recently signed a bill to procure a new fleet of refueling aircraft, and in October 2006, the Air Force stated its goal of procuring 450 converted Boeing 767s (21); however, military procurement is a notoriously slow and uncertain proposition. While the need for tankers is not diminishing and may increase over time, the future of any proposed increase to the service or ability of the current fleet remains uncertain. The one certainty is that at this time the United States owns a limited fleet of tankers which must be utilized to the best of their capability. Therefore, to gain future capability from the current fleet the methods of planning must be optimized.

Figure 10: 767 Refueling

2 *Problem Description*

In 2006, the United States Air Force Office of Scientific Research (AFOSR) approached CASTLE Laboratory at Princeton University to develop an aerial refueling simulator. The proposed simulator was required to model and plan aerial refueling operations, as well as answer the myriad of questions about optimal tanker placement, tanker deployment, and optimal receiver refueling. To aid the development of an aerial refueling model, the current Excel mission planning program in use at AFOSR was given to CASTLELAB. In the current Air Force model, an operational planner specifies the type of planes requiring refueling, when the planes need refueling, and where they will require refueling (refueling locations are referenced as tracks). Given those inputs, the Air Force model sequentially determines the receiver requirements and assigns a tanker to a receiver at the receiver's assigned track. Within the AFOSR model the refueling tracks are given as inputs. When assigning a tanker to a receiver the model first determines if a tanker is already at the track and attractive to refuel the receiver. If the tanker is currently refueling a receiver or low on fuel another tanker is assigned to the receiver. The model uses a myopic policy exclusively and does not examine any future values of holding tankers at a track. Therefore, while the model is an adequate planning tool it does very little to approach the goal of optimizing tanker usage.

Given the current AFOSR model, and the requirements that a future aerial refueling model both plan and optimize, the proposed simulator provided a perfect use for Approximate Dynamic Programming. Using ADP a simulation package was created which simulates and optimizes receiver and tanker movements. The current AFOSR model has the receivers refueling tracks and times as given inputs which limits any optimization in the system strictly to the movements of the tankers. While optimizing tanker movements is not a trivial exercise, it can be accomplished through standard simulation and does not create much value for the mission planners. In CASTLELAB the problem was approached in a more holistic manner, removing fixed receiver refueling tracks such that both the tanker and receiver movements are optimized within the system.

Since the CASTLELAB model removes the refueling tracks as a constraint in the system, a proxy for refueling location was required to guarantee receiver mission success. The aerial refueling model uses the refueling time as the hard constraint to determine "when" the mission will be refueled; however, it is left to the model to determine "where" the receiver will be refueled. The approach used in CASTLELAB allows for receiver and tanker movements which optimizes fuel usage by both entities. While the model solves the optimal placements of tankers and receivers it does not relegate the central goals of the receiver missions: arriving to a target at a specific time and with a specific fuel load. These constraints are hard coded in the AFOSR model but in the aerial refueling model they are used as soft constraints which guide the movements of receivers in the system. By eliminating the hard constraint and replacing it with a soft constraint it allows the model to optimize behavior while also fulfilling the receiver mission goals. Also built into the model are tunable parameters which can further refine receiver movements(ie favoring shorter refueling track to target movement).

The approach taken in CASTLELAB is general in nature yet specific in practice. This allows for the use of proven optimization algorithms and problem specific requirements. Throughout the thesis, refinements of the model are discussed and further possible extensions posed. The model and results shown in the following sections are powerful demonstration of how ADP is used for planning the refueling of the US military in the future.

2.1 *Approximate Dynamic Programming Method*

The aerial refueling problem is formulated as a multi stage model in which decisions are made sequentially. The problem was approached as a resource allocation problem which could be solved using Approximate Dynamic Programming (ADP). ADP is an extension of Dynamic Programming and Bellman's equation; however, while dynamic programming requires the enumeration of every state to solve Bellman's equation (usually impossible), ADP is an iterative simulation strategy which does not require the enumeration of all states. During each iteration of a simulation, decisions are made using knowledge gained from previous iterations and after each decision information about the state of the system is acquired. The information collected in the form of marginal cost and value functions is then incorporated with the previous knowledge of the system, and the accumulated knowledge is used to make decisions in the next iteration. Therefore, every decision "sees" all previous knowledge of the system and attempts to minimize(maximize) the cost of the decision to find the optimal solution.

The specifications of the model and how information is gathered and incorporated are described in great detail for both the general ADP framework and the aerial refueling model. The description of the ADP framework follows the guidelines set forth in Warren Powell's forthcoming Approximate Dynamic Programming text (17). The following sections highlight the specifications of modeling in ADP and the algorithmic strategy used in creating the aerial refueling model. Topics discussed include: modeling resources, the decision variables and functions, the measurement of the state of the system, how the information process in structured, the transition of resources within the model, a general overview of policies guiding model behavior, and how the system is measured at a single point in time which includes the objective function of the model.

2.2 *Why Not Dynamic Programming or Linear Programming?*

When looking at the optimal assignment of tankers to receivers from a perspective of 10,000 feet, the approaches of linear programming or dynamic programming appear to be reasonable methods to solve the aerial refueling problem. Using a linear programming formulation, a series of sequential networks with receivers acting as the demands and the tankers providing the supply nodes could be set up. This scheduling approach is used in Chemical Engineering where different processes occur in time and one reaction ending must coincide with the beginning of the following process. However, upon coming down from the high view and drilling into the actual demands of the problem, the shortcomings of the network approach are obvious. Using linear programming the assignment of two tankers and two receivers to two tracks is not a daunting task on the surface. However, the complexities of the system inherent to nonlinear cost which are not readily apparent make solving the problem much more difficult.

When refueling receivers, the cost associated with refueling two receivers by a single tanker is different than having each receiver getting refueled by their own tanker. This is due to the cost associated with queuing which can occur in a simulation and must be incorporated into the overall cost. Therefore, for this simple problem the cost of having a different tanker for each receiver as well as the cost of having two receivers assigned to one tanker must be explicitly calculated. Additionally, the cost of moving the tankers to and from each track, and the cost of moving the receivers to each track and then to their target all have to be calculated to obtain the cost of having tankers and receivers at various tracks.

In this small example if the two tankers and two receivers are identical then the permutations of the cost can be calculated, but if the tankers as well as the receivers are different then the problem becomes increasingly complex as multiple simulations would be required. Also, many constraints to the system such as maximum queuing time per receiver and refueling rates for each tanker/receiver combination must somehow be incorporated. When examining the problem at a lower level it becomes apparent that a network approach is not feasible to solve the problem with all of its built-in complexities. While alternative approaches such as branch and bound strategies could be implemented, there is not a simple linear programming approach.

The examination of dynamic programming is very similar to that of linear programming in that when viewed from a high level it appears to be a reasonable approach. The shortcomings come in very quickly with a phrase familiar to individuals versed in dynamic programming: "the curse of dimensionality". For those unversed in dynamic programming the following explanation of the curse will quickly make apparent why a strict dynamic programming solution is not feasible.

If an individual is standing on a street corner, and will flip a coin twice to determine if he will go north one block, east one block, west one black, or south one block, a transition matrix for the location of the individual in the next period can easily be determined. After the first period the individual flips the same coin again and makes the same decision. Again a transition matrix could be used to determine the probabilities of the man's final location. After the second period the individual could be in any of 9 different positions as shown in Figure 11:

Making the assumption that ending up at each location has a path dependent cost associated with it such that moving east/west does not have the same cost as moving west/east despite ending at the same location. This is a reasonable assumption given the following example: If the individual is at the top of a hill when at the center position and they move east with their first move they move down the hill; however if for their first move they move west they remain on flat terrain illustrated in Figure 12 then there are 9 locations possible and 16 costs associated with the two moves

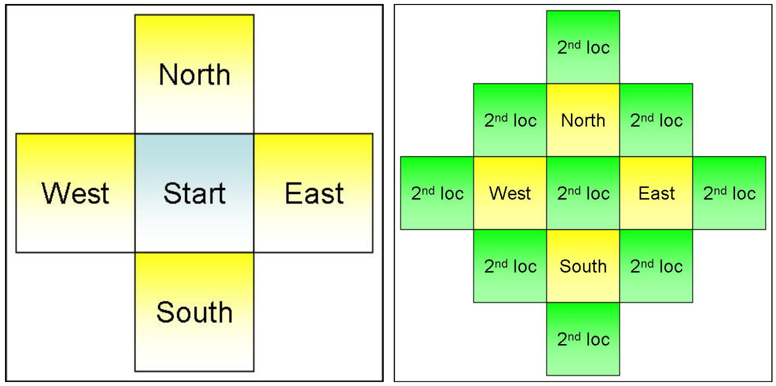

Figure 11: Locations for One Stage and Two Stage Move

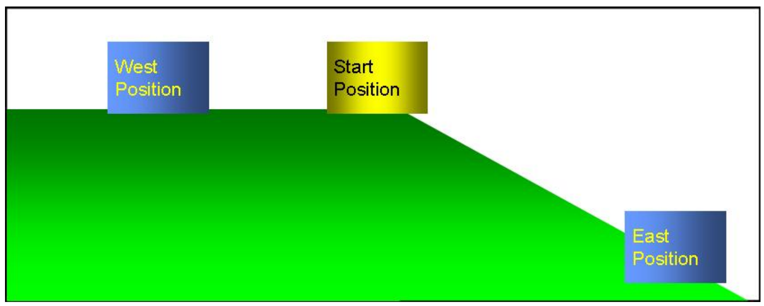

Figure 12: Example of Path Dependence

(Cost shown by Equation 1).

cost matrix= (1)

(1stmove/2ndmove)	north	south	east	west
north	nn	ns	ne	nw
south	sn	ss	se	sw
east	en	es	ee	ew
west	wn	ws	we	ww

To measure the system and determine the state of the system after two moves, the 16 costs associated with the moves are required but not the 9 locations which are implicitly given in the cost. If the example was extended to include more realism, such as knowing if the individual moving is a man or women as well as their age, then to measure the system those factors would have to be included. Including that the individual could be a man or a women as well as any of 50 ages, the space which could possibly be reached and must be enumerated grows to $16(movement - cost) * 50(ages) * 2(sexes) = 1600(states)$. As shown in this brief example is easy for the state of the system to become incredibly large by adding complexity to the system, and thus dynamic programming methods get bogged down for all but the smallest problems. In the aerial refueling problem the complexities far outstrip the given example and it would be computationally intractable to enumerate all the states of the system. Therefore, while dynamic programming provides the backbone for the problem it cannot be used directly.

2.3 *Bellman's Equation - The Foundation*

The foundation of ADP lies with a series of dynamic programming equations known as Bellman's equations:

$$V_t(S_t) = max_{x_t \in \mathcal{X}_t} \mathbb{E} \left\{ C_{t+1}(S_t, x_t) + \gamma V_{t+1}(S_{t+1}) \mid S_t \right\}. \tag{2}$$

Bellman's equations focus upon making decisions, x_t, at a distinct time epochs using both the immediate associated cost of the decision, $C_{t+1}(S_t, x_t)$, and any future value associated with that decision, $\gamma V_{t+1}(S_{t+1})$. Within Bellman's equation is the idea of the "state" of the system, S_t, which is used to compute both current and future values. A "state" as defined by Powell as

> "the minimally dimensioned function of history that is necessary and sufficient to compute the transition function, contribution function and the decision function." (17).

For the aerial refueling model, the state of the system includes all the information about the tankers and receivers in the system at a given point in time. At time t the state of the system is measure of where tankers and receivers are located, the fuel levels/demands of the tankers/receivers, the refueling times associated with the receivers as well as any currently occurring movements of the tankers in the system. The state of the aerial refueling model is an all encompassing variable which provides the knowledge of what is happening throughout the system.

Bellman's equation, while elegant, suffers from the three curses of dimensionality which limit its usefulness in practice. The state space is the first curse since even in small problems with few resources the state space grows exponentially with the addition of more resources. The state space has dimensionality of $|\mathcal{A}|$ which in the aerial refueling model is a combination of all the attributes of a tanker. The attributes of the tanker which are further refined in section 2.4 include the tanker's fuel level, location, base, id number, and other important aspects of the tanker required in the model. The second curse is the action space which incorporates the decision sets of the system, $x_t \in \mathcal{X}$, as well as the state space. The action space is a function of both the state space $\mathcal{A}$ and the decision space $\mathcal{D}$ (The decision space is the set of all decisions possible). The size of the action space is a vector of dimension $|\mathcal{A}| * |\mathcal{D}|$ which is incredibly large in all but the smallest of problems. The last curse is the outcome space which is $|\mathcal{A}| + |\mathcal{B}|$ dimensioned where $\mathcal{B}$ is defined as the information space.

While solving dynamic programs using Bellman's equation proves intractable for all but the smallest problems, through manipulation the equation provides the basis for solving problems using ADP. One of the major hurdles in solving Bellman's equation is the expectation, $\mathbb{E}\left\{C_{t+1}(S_t, x_t) + \gamma V_{t+1}(S_{t+1}) \mid S_t\right\}$, which cannot be solved except for small deterministic problems! To solve Bellman's equation a recursive strategy is used which eliminates the expectation and uses sample realizations (17). As a primer for approaching the following series of equations, those unfamiliar with pre and post decision states, resource states, or value functions should skip to the next several sections 2.4-2.6 where they are described.

Solving the optimal policy in Bellman's equation is done by breaking the equation into two steps and applying a recursive strategy. The two steps of the recursion are set up as follows:

$$V_{t-1}^x(S_{t-1}^x) = \mathbb{E}\left\{V_t(R^{M,W}(S_t, W_{t+1})) \mid S_{t-1}^x\right\} \tag{3}$$

$$V_t(S_t) = \max_{x_t \in \mathcal{X}_t} \left[C(S_t, x_t) + \gamma V_t^x(R^{M,x}(S_t, x_t)) \right]. \tag{4}$$

The second equation is substituted into the first which produces:

$$V_{t-1}^x(S_{t-1}^x) = \mathbb{E}\left\{ \max_{x_t \in \mathcal{X}_t} [C_t(S_t, x_t) + \gamma V_t^x(S_t^x)] | S_{t-1}] \right\} \tag{5}$$

In equation 5 the post decision state variable is used and therefore the expectation can be dropped. The last equation can then be solved using a sample realization $W_{t+1}(\omega)$ from $\omega \in \Omega$. In this manner the value function $V_t^x(S_t^x)$ is replaced with an approximate value $\bar{V}_t(S_t^x)$ from a single sample . The decision function can then be set up and solved:

$$X_t^\pi(S_t) = \arg\max_{x_t \in \mathcal{X}} \left[C_t(S_t, x_t) + \gamma \bar{V}_{t+1}(S_t^x) \right]. \tag{6}$$

The decision, x_t^n, is identified both for the time period in which is occurs, t, as well as the iteration, n, of the algorithm. In a large model it is reasonable to take a monte carlo sample to create the sample path from a space of possible outcomes. However, within the aerial refueling model the sample path is the receiver missions, which are established prior to the start of the simulation and followed while stepping through time. In solving the decision function above at iteration n the approximation of a value function of the state from a previous iteration is used instead of the expectation of a future state. Therefore, through replacing $\bar{V}_t(S_t^x)$ with $\bar{V}_t^{n-1}(S_t^x)$, where $n-1$ denotes value function approximation from the previous iteration, the equation can be explicitly solved.

2.4 *The Attribute Space of Aerial Refueling*

The attributes of the model are important in explaining its evolution and its current state. The vocabulary of dynamic resource management is used throughout the model description (17). Within this framework tankers are "resources" and receivers are "tasks". The attribute vector, a, defines the state of a single tanker resource.

The tankers are defined by a collection of attributes which are both numerical and categorical.

$$
\begin{aligned}
a &= \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \\ a_7 \end{pmatrix} = \begin{pmatrix} \text{Location} \\ \text{Base} \\ \text{Fuel Level} \\ \text{Tanker Type} \\ \text{Usage} \\ \text{ID} \\ \text{BeenUsed} \end{pmatrix} \in \mathcal{A} \\
\mathcal{A} &= \text{Set of all possible tanker attributes } a.
\end{aligned}
$$

The categorical attributes such as Tanker Type and Location are easy to enumerate since they come from a predefined set. However, for a continuous attribute such as fuel level it is not possible to enumerate all values. The attribute space of a tanker is used to define the value of a tanker, and it is incredibly difficult if not impossible to value a continuous attribute space. As an example, for a tanker at a refueling track, is it important to make a distinction between a tanker having 100,000 lb of fuel or 105,000 lb? The answer for the model is no, it does not matter for such a small difference, but if the difference were 50,000 lb of fuel then there could be quite a large difference in the value of the tanker. While the attribute space is defined as continuous, when the values of tankers are computed the continuous attributes are discretized and the continuous attribute space becomes a discrete attribute space. The set spanning all possible attribute spaces is referenced as $\mathcal{A}$.

The receivers “tasks” also have attributes vectors:

$$
\begin{aligned}
b &= \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \\ b_6 \\ b_7 \\ b_8 \end{pmatrix} = \begin{pmatrix} \text{Type} \\ \text{Track Arrival Time} \\ \text{Track Exit Time} \\ \text{Mission Number} \\ \text{Type} \\ \text{Base} \\ \text{Offload} \\ \text{Target} \end{pmatrix} \in \mathcal{B} \\
\mathcal{B} &= \text{Set of all possible receiver attributes } b. \qquad .
\end{aligned}
$$

However, in the model the receiver attribute space, B, is not used to estimate the value of the system being in a state. While it is possible to estimate the value of having a receiver in the system, the model subordinates the receiver movements to the tanker movements. The value of a receiver in the system is conditional upon the tanker movements. The receiver movements in the system are guided through a policy which uses the location of tankers in the system. When there are multiple tankers at different tracks, each receiver is assigned to the track which minimizes its individual distance to the track and subsequent movement to its target. Since the tanker locations and quantities determine where receivers move, due to their policy, the receiver refueling cost is captured in the value functions of the tankers.

2.4.1 *Aggregating the Attribute Space*

The tanker attribute space holds all relevant information about each tanker in the system; however, it is cumbersome to compute the value of each tanker using all information from the attribute vector. When computing the value of a tanker at a track, it is obvious that knowing the fuel level is important, but does knowing the tanker ID have any value? In this model the answer is "no" for two reasons. The first reason is that the specific ID does not provide any actionable information for the system. Knowing the ID of the tanker does not tell the system if the tanker is low on fuel or if it can refuel a specific type of receiver. The tanker ID is extraneous information when making a decision in the system since it has no impact on the value a tanker can provide in the system.

The second reason using the tanker ID does not benefit the system is that value functions created using the tanker ID are too narrowly defined within the system. If a value function is identified by the tanker ID number, then that value function is only representative of the value of that specific tanker. Obviously when creating value functions they should be specific enough to provide actionable information but general enough so that they can be applied to multiple similar tankers.

Therefore, a value function which uses fuel level is appropriate but a function which uses tanker ID is not. Using the fuel level in a value function provides knowledge to the system since the value function is applicable to all tankers at the time point with a similar fuel level. While different algorithmic strategies can implement more or different attributes in determining the value of a tanker, the general form can be thought of as taking an attribute space, a, and simplifying it when calculating values of the attribute space. The aggregation function takes a very detailed attribute space and simplifies it to a more tractable and usable form.

$$G^g : A \rightarrow A^{(g)} \tag{7}$$

The function above is the aggregation function where $A^{(g)}$ represents the g^{th} level of aggregation of attribute space A. For approximating the value of an individual tanker in the model the aggregation function $a^{(2)}$ used was:

$$a^{(2)} = \begin{pmatrix} Location \\ FuelLevel \end{pmatrix} \tag{8}$$

While it appears that a lot of information was lost due to aggregation, the information still exists attached to each tanker. Within the model the attributes such as base location and tanker ID are not discarded; however, when valuing a tanker the extraneous information is parsed out so that the value function can be extended to nearly identical tankers.

2.4.2 *Extending the Attribute Space to the Resource State and Time*

When modeling time, the attribute vector, a, is indexed by the time period in the system, t. The notation a_t identifies the attribute of a single tanker at the time t. Extending the single tanker example up to the multiple tanker realities of the system requires the introduction of the resource state variable. When multiple tankers have identical attributes, the resource state captures the tankers as follows:

$R_{ta} =$ The number of resources with attribute vector a at time t.

$R_t = R_{ta \in \mathcal{A}}$ The collection of all resources, $\mathcal{A}$ is the entire attribute space.

R_t is known as the *resource state vector*.

2.4.3 *Pre and Post Decision Resource State*

The aerial refueling simulation occurs in continuous time; however, to model the system it is broken into discrete time intervals. The discrete time intervals allow for the notion of the resource state in reference to the decision epochs. At a decision epoch, the decisions, x_t, about the tanker movements in the next time period are made. After the decision, exogenous and endogenous information about the system is collected in the information state, W_t. The progression of the history process is defined as:

$$h_T = (R_0, x_0, W_1, R_1, x_1, W_2, ...R_{T-1}, x_{T-1}, W_T, R_T)$$

The above formulation is a natural way to make a decision, collect information, evaluate the current state, and make the next decision. Within this formulation the resource state, R_t, is defined as the pre-decision resource state. In the aerial refueling problem, the pre-decision resource state is used to determine the locations of tankers and the available actions for the tankers. The aerial refueling problem has the added complexity of receiver queuing and refueling, and the pre-decision resource state cannot guide receiver policy movements. If the system only had a pre-decision resource state, then two decisions about moving tankers and moving/refueling receivers would have to be made simultaneously. The problem would get very messy since it would face the impossible task of deciding where to send receivers before the movements and locations of the tankers are known. To resolve this quagmire, the post decision resource state, R_t^x, is used as shown in the following history process.

$$h_T = (R_0, x_0, R_0^x, W_1, R_1, x_1, R_1^x, W_2, ...R_{T-1}, x_{T-1}, R_{T-1}^x, W_T, R_T)$$

The post decision resource state, R_t^x, "sees" all the information of the pre decision resource state and the decision x_t. For the refueling model this simplifies the decision making process for the tankers and subsequently the decisions about the receiver movements. At time t the tanker movement decisions are made, which transforms the resource state vector from R_t to R_t^x. Within R_t^x is the explicit knowledge of the actions during the following time period $t+1$, and the post decision state variable provides actionable information to the system. When R_t^x is known, the actions of the tankers during time period $t+1$ are knowledge to the system. Tankers at tracks which are being held at a track for period $t+1$ are seen by receivers arriving to the system during time interval $(t, t+1]$. The receiver movement decisions policy guides the receivers to tracks with tankers and the problem of simultaneous decisions making disappears. The pre and post decision states will be used throughout the rest of this thesis, with the post decision always denoted by superscript, x.

2.5 *The State Variable*

As defined earlier, the state variable holds the information necessary to compute the transition, objective, and decision functions. The state variable at time t is defined as S_t, but what is contained in S_t? In the general framework of ADP the state vector is a composite of the resource state and the demand state, D_t. The demand state is the state of all the receiver missions entering the system at time t.

$$S_t = (R_t, D_t).$$

Once again the aerial refueling model has the added complication that decisions are not made solely at decision epochs, but also within time periods. This leads to the complication of when to measure the state variable. For the sake of clarity, the state variable will always be measured at the decision epoch. Another complication of the model is that demands do not disappear if they are not satisfied. The unsatisfied demands from previous time steps remain in the system until they are satisfied (ie receivers will not simply disappear if they are not refueled in a single time period). To illustrate the process which is used in the aerial refueling model, the history process below clarifies when the state variable is measured:

$$h_t = (S_0, x_0, S_0^x, W_1, S_1, x_1, S_1^x, W_2, S_3......S_{t-1}, x_{t-1}, S_{t-1}^x, W_t, S_t).$$

Within the history process, S_t is measured just before decisions are made in the model and sees both the resources and the remaining demands from previous time periods. The state variable must see the remaining demands so that a decision to move a tanker to base is not made when a receiver is currently waiting in queue. The model uses the state variable to make the decisions, x_t, about moving tankers. After the decisions have been made the demands of the receivers entering the system during time period $(t, t+1]$ become known to the system. As the receivers arrival to the system become known, a second set of decisions is made about receiver movements. In the history process, the exogenous information process W_{t+1} is a measure of two exogenous information processes: the update of the attributes of the tanker (ie fuel level), and new receivers entering the system.

$\hat{R}_{ta}$ = The change in the number of tankers with attribute a due to information arriving during time interval t. Within time period t the tankers can be in use, refueling, or recently released from fueling a receiver.

$\hat{D}_{tb}$ = The change to the receiver missions with attribute b during time period t due to refueling or entering a queue.

Within the system $W_t = (\hat{R}_t, \hat{D}_t)$ is used as the generic variable for new information that arrives in time period t. Implicit to the information process for the aerial refueling problem are the receiver movements which are guided by a policy which uses S_t^x. Additional new information within W_t is a tanker/receiver fuel level alteration, tankers moving from a previous time period reaching its location, or receivers entering a queue and being assigned new expected refueling times $t' > t$.

Therefore, in the aerial refueling model the state variable is not simply the resource and demand state at time t. Rather it is a composite of the resource state at time t

and the demand vector from time period t as well as the information process of the system.

$$\begin{aligned} S_t &= (R_t, D_t) \\ &= S^M(S_{t-1}, x_{t-1}, W_t) \\ &= S^{M,W}(S^x_{t-1}, W_t) \end{aligned}$$

2.6 *The Decision Sets*

In a traditional resource allocation problem there is a single layer of decisions which are made at decision epochs. However, as alluded to previously when discussing the state variable in the aerial refueling model, the decision process for each period consists of sequential decisions. The first decision concerns the movement of tankers and the subsequent decision the receiver movements. For the aerial refueling model, the first set of decisions at the decision epoch create the second decision set and are therefore more important. Additionally, the first set of decisions are formulated as a linear programming network at each time period which use the value functions to make decisions as guided by Bellman's equation. The decisions for the tankers are set up as follows:

$d =$ An elementary decision which will act upon a resource (Moving or Holding a Tanker)

$\mathcal{D} =$ The set of all possible decisions. (Move Tanker to Track, Hold Tanker At Track, Move Tanker to Base, Hold Tanker at Base)

$\mathcal{D}_a =$ The set of all possible decisions that can act on a resource with attribute a.

The composition of $\mathcal{D}_a$ is defined by the location of a tanker and whether it is refueling a receiver at the decision epoch. Tankers that are currently refueling a receiver are not allowed to stop refueling to make a separate decision but rather will complete refueling and have the singular decision of Hold Tanker At Track. Also, a tanker at a track does not have the decision to move to an adjacent track, but rather its decision set consists of holding at the current track or returning to its base. Further refining the model and the decision sets:

x_{tad} = The number of times decision d is applied to resource with attribute vector a. In the aerial refueling problem there are often several tankers with identical attribute vectors such as a KC-135, with full fuel, at base available for use.

$x_t = (x_{tad})_{a \in \mathcal{A}, d \in \mathcal{D}}$

$\mathcal{X}_t$ = The set of all possible actions, x_t, at time t

At each time period the model is set up as a myopic linear program shown in Figure 13, which produces the following constraints:

$$\begin{aligned} \sum_{d \in \mathcal{D}} x_{tad} &= R_{ta} & \forall a \in \mathcal{A}, \\ \sum_{d \in \mathcal{D}} x_{tad} &= l_{tad} & \\ x_{tad} &\geq 0 & a \in \mathcal{A}, d \in \mathcal{D}. \end{aligned}$$

The first equation is the flow conservation constraint which guarantees there are equal tanker decisions and tankers available. The second equation guarantees that there are not more decisions made than a specified upper limit l_{tad}. X_t is the set of all feasible solutions x_t to the above constraints. The decisions x_t are determined by a decision function.

While Figure 13 shows the general network of the tanker movements, it leaves out a very important aspect: why would the tankers move? Figure 14 introduces value function approximations which help to explain what the linear program is trying to maximize and why tankers move.

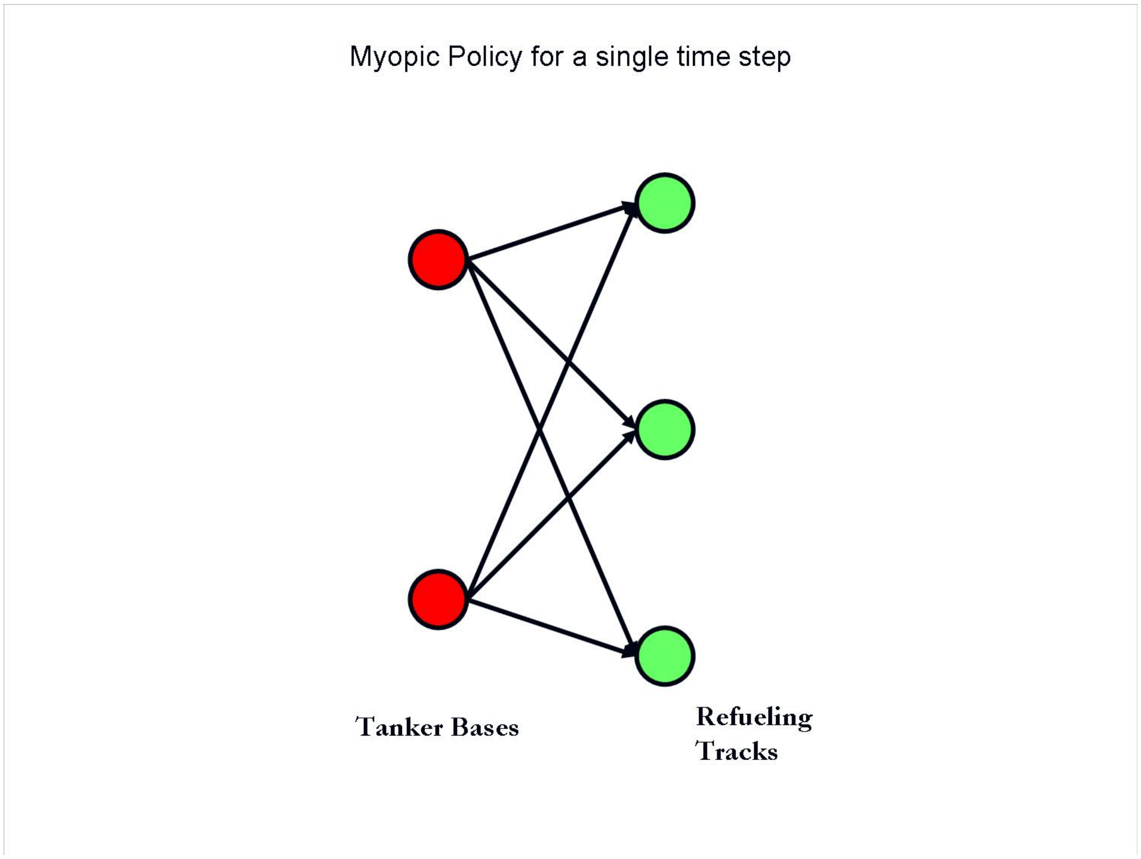

Figure 13: Myopic linear program

When a tanker moves from its base to a track, it accrues a negative cost (fuel burned); however, there are rewards for a tanker at a track such as refueling receivers which would otherwise fall from the sky. At each refueling track node there are associated value function approximations which represent the positive values of refueling a receiver at that track. The value function approximations will be discussed further in section 2.9; however, it is easy to think that each arc of the value function represents the positive value of refueling a receiver or group of receivers with varying numbers of tankers.

2.6.1 The Receiver Policy Decisions

The receiver movements within the system are guided through a decision policy. The receiver decisions occur after the decision epoch and are dependent on the tanker

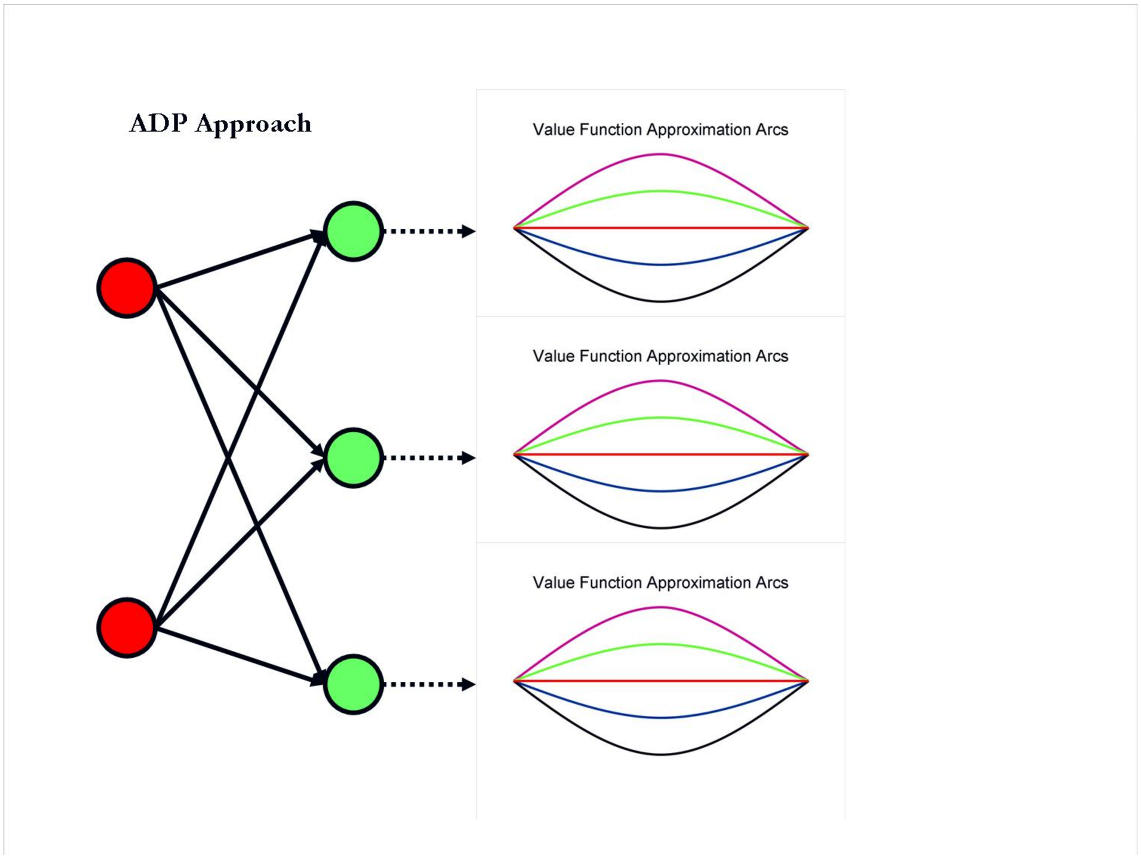

Figure 14: Myopic linear program with value functions

decisions, x_t. After the tanker decisions have been made, the receiver demands are introduced into the system:

$$D_t = (D_{tb})_{\in \mathcal{B}} = \text{The set of all receiver demands.}$$

$$D_{tb} = \text{The number of receiver missions of type } b.$$

When the demands are introduced, the decision set for the receivers is created through a predetermined policy. While the tanker decision set is solved through a linear program, the second decision set is constructed through a previously created policy function. The policy is constructed such that the receivers entering the system must move to the set of available tracks which have tankers while minimizing total distance traveled.

$\mathcal{Y} =$ The set of tracks.

$y \in \mathcal{Y} =$ Particular track.

$c_{by} =$ Cost of assigning receiver with attributes "b" to track "y".

The set of all tracks is further divided into tracks which currently have a tanker. Receivers cannot be assigned to tracks without tankers. Therefore, the set of all the tracks is looped over to find the subset with tankers.

$\mathcal{Y}' \subset \mathcal{Y} =$ Subset of all tracks which currently have a tanker.

$\mathcal{Y}' = \sum_{y \in \mathcal{Y}} y * 1_{tanker,y}$

If the subset of tracks with tankers is empty then the receiver missions are recorded as failures in the system. If the subset is not empty, the receivers are assigned to the track which has the lowest associated cost.

$y_r =$ Track chosen for receiver r.

$y_r = \arg\min_{y \in \mathcal{Y}'} C_{by}$

Once the receivers have been assigned to their respective tracks, the model sequentially assigns them to the available (not currently refueling) tankers at the track. If all tankers at a track are refueling other receivers then the model sequentially assigns the receivers to the queues of the refueling tankers.

2.7 *Transition Function*

During the simulation both the resources, R_t, and demands, D_t, evolve over time. The evolution of the demand focuses on the assignment of receivers to tracks and their refueling. The resource vector, R_t, evolves from endogenous and exogenous factors. The first factor in resource state evolution is due to decisions (Move Tanker to Base, Hold Tanker on Track..). The resource state after a decision has been made is called the post decision resource state R_t^x. This post decision resource state is an important aspect of the model since it determines the availability of the tankers to refuel receivers at a track. Endogenous information about the resource state, $\hat{R}_t$, arrives to the system in the time period $t-1$ to t. An endogenous information process occurring in the model is the depletion of fuel from a tanker when it is refueling a receiver. There are also exogenous events that effect the resource state; however, their notation varies slightly.

To illustrate the evolving states of the system, a single tanker at the attribute level will be used. At time $t = 10$, a tanker with attribute vector a_{10} (which will be limited to the tanker's available time, and location) has been assigned the decision to hold at its track until $t = 20$. The post decision attribute a_{10}^x has two consequences for the system. The first is that the tanker is expected to be available for a new decision at $t = 20$, and the second in this multi-stage process is that the tanker is available for refueling assignments immediately at $t = 10$ until $t = (20 - \epsilon)$. If a receiver enters the track at $t = 18$ and is assigned to the tanker then the tanker is now "in use" refueling the receiver. Assuming that the tanker takes five time units to refuel the receiver, the new information has changed the attribute vector of the tanker, $\hat{a}_{18}$. When the decision epoch at $t = 20$ is reached, the tanker no longer has the attribute vector from a_{10}^x but rather a transformed attribute vector. The tankers pre decision attribute vector a_{20} now has the tanker available at time $t = 23$.

The first change in the attribute vector(hold at track which determines the tankers availability time) is a result of the decision made at the epoch. The second change in the attributes occurs due to new information arriving (the assignment of the receiver to the tanker and the receivers refueling). The first change is represented in the model using the function:

$$a_t^x = a^{M,x}(a_t, d)$$

The effect of the new information on the system is represented by the function:

$$a_{t+1} = a^{M,W}(a_t^x, W_{t+1}).$$

In the second function the term W_{t+1} represents the new information arriving to the system in the time period from t to $t+1$. The functions $a_t^x = a^{M,x}(a_t, d)$ and $a_{t+1} = a^{M,W}(a_t^x, W_{t+1})$ show the physics and the decision making rules of the system. If a decision acts on the tanker with attribute a_t, then $a^{M,x}(a_t, d)$ determines if a tanker will be available to refuel receivers. As a continuation of the previous example, the post decision attribute a_{20}^x has the tanker staying at the track and available at time $t = 23$; therefore, $a^{M,x}(a_{20}, d)$ knows when the tanker is available and when it will be available for its next movement, $t = 30$.

Extending the attribute vector to the full resource vector, the first transition function process is:

$$R_t^x = R^{M,x}(R_t, x_t).$$

The second transition function process is represented by:

$$R_{t+1} = R^{M,W}(R_t^x, W_{t+1}).$$

However, in practice the resource vector is often written as a transition equation, $R_{t+1} = R^M(R_t, x_t, W_{t+1})$. Within this model indicator functions are used to facilitate the ease of movement between the modeling and algebraic realities of solving the problem. The indicator functions below use the notation of a' as the post decision attribute vector :

$$\delta_{a'}^x(a_t, d) = \begin{cases} 1, & \text{if } a^{M,x}(a_t, d) = a' \\ 0, & \text{otherwise} \end{cases}$$

$$\delta_{a'}^W(a_t, W_{t+1}(\omega)) = \begin{cases} 1, & \text{if } a^{M,W}(a_t, W_{t+1}(\omega)) = a' \\ 0, & \text{otherwise} \end{cases}$$

The post decision transition $R^{M,x}(R_t, x_t)$ function is given by:

$$R^{M,x}_{ta'} = \sum_{a \in \mathcal{A}} \sum_{d \in \mathcal{D}} \delta^x_{a'}(a_t, d) x_{tad}$$

The transition function $R^{M,X}(R^x_t, \hat{R}_{t+1})$ is given by the post decision state variable and the exogenous information process that changes the state variable:

$$R_{t+1,a} = R^x_{ta} + \hat{R}_{t+1,a}$$

Within the model the transition function for the demands plays an important role and is similar in structure to the resource state. The demand state variable, D_t, can be represented in two stages with state dependent decisions. The effect of decision d on a receiver with attribute b_t can be represented using functions $b^{M,x}(b_t, d)$ and $b^{M,W}(b^x_t, W_{t+1})$, which correspond to $a^{M,x}(a_t, d)$ and $a^{M,X}(a^x_t, W_{t+1})$ in the dynamics of the system; however, the decisions are from different sets. As a receiver with attributes b_t arrives in the system, and a decision d_t is made to send the receiver to a track, the receiver is transformed to b^x_t. The vector b^x_t now has the track and the refueling time. At time $t' \geq t$, the receiver arrives at its track and is assigned to a tanker. However, at this point the receiver can enter into a queue and change the refueling time to $t + \epsilon$. Such transitions occur frequently in the model and its important to realize that both the receivers and the tankers evolve over time.

2.8 *The Contribution Function*

The objective of this model is to minimize total fuel usage by both tankers and receivers. Since this problem is a two stage process, there is the added complexity that the second stage contribution function depends on the outcome of the first stage. The general model for a two stage contribution function is of the form:

$$C_t = C_{t,1}(x_{t,1}) + \mathbb{E}C_{t,2}(x_{t,2}) \tag{9}$$

Within equation 9 the first and second stage decisions as well as the first and second stage contributions are denoted by a subscript 1 or 2. For the aerial refueling model the contribution of the first stage is the cost of moving or holding a tanker which is a known value. The function for calculating the first stage is shown below and uses the form such that c_{0ad} is the contribution of making a decision d on a tanker with attribute a in the first stage:.

$$C_{t,1}(x_t) = \sum_{a \in \mathcal{A}} \sum_{d \in \mathcal{D}} c_{tad} x_{tad} \tag{10}$$

The contributions for stage one are deterministic (hold tanker, move tanker) and are calculated as a function of time spent in the air. Within the aerial refueling model the second stage contribution function is determined by the first stage decisions. Additionally, the second stage contribution function for the refueling problem is not linear or deterministic, but rather must be explicitly calculated through simulation. The reason for the non-linearity is the queuing within the system. The contribution of assigning receivers to tankers for refueling cannot be assumed to be linear since as the queue grows in length, the contribution of assigning an additional tanker grows in a piecewise manner. The first receiver assigned to a tanker immediately begins refueling and the contribution is linear with respect to fuel required and refueling rate. If the next receiver added to the system arrives while the first receiver is refueling, then it is added to the queue and must wait behind the first receiver before refueling at the tanker. This process is repeated for every additional receiver added to the queue. When a queue accumulates from an unfulfilled receiver mission, D_t, and the incoming receivers, D_{t+1}, the queue must be simulated to find the contribution. Figures 15 and 16 illustrate the queuing problem. The table has a single time period. At the beginning of the period Receivers 1 and 2 are in the queue and Receivers 3 and 4 join the queue in at different points in $t+1$. These figures illustrate that the second stage contribution during time period $t+1$ is both a function of refueling and queuing times, and is dependent on the number of tankers at a track. They also show how receivers entering during time period $t+1$ can make a second stage contribution to $t+1$ as well as later time periods, as is the case with receiver four.

The second stage contribution function cannot be written in similar fashion as

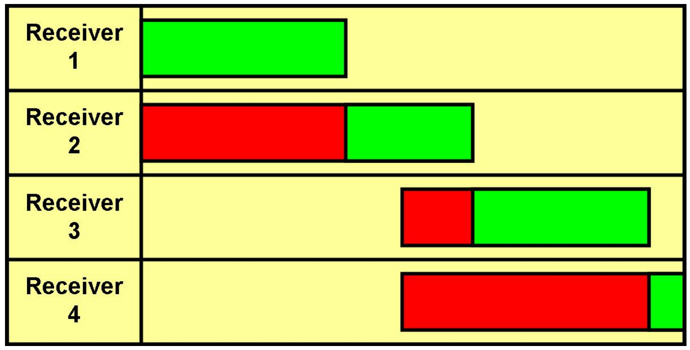

Figure 15: Refueling Receivers with 1 Tanker at a Track: Refueling (Green) - Queuing (Red)

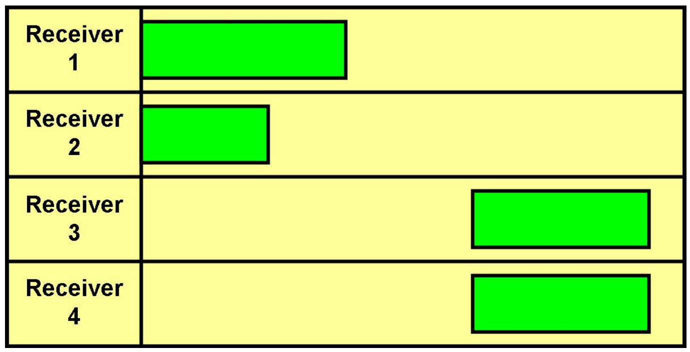

Figure 16: Refueling Receivers with 2 Tankers at Track: Refueling (Green)

the first stage due to the non linearity of the queuing cost. A more representative function for the second stage contribution is formed by replacing the expectation in Equation 9 with the explicit cost of the queuing and refueling cost. The cost of the queues is a scalar value added to the contribution of the decisions x_t. The value is a function of the post decision resource state and the receiver demands represented by:

$Q(R_t^x, D_{t+1})$ = The explicit cost of refueling receivers.

The total contribution for decisions x_t and period $t+1$ is therefore a combination of the tanker movement cost and the receiver refueling and queuing cost:

$$C_t(R_t, x_t) = \sum_{a \in \mathcal{A}, d' \in \mathcal{D}} c_{tad} x_{tad} + Q(R_t^x, D_{t+1}), \tag{11}$$

Q is a function of the post decision resource state (the tankers holding at a track or in use from the previous period) and the demand state (the queue to which the receivers are assigned). In the aerial refueling model the two stages are calculated separately. The first stage is calculated at the decision epoch t, and the second stage is computed during the time interval $t+1$ through simulation.

While the second stage of the contribution function can be calculated through simulation, it has the shortcoming at time t in that it cannot see the value of $Q(R_t^x, D_{t+1})$, and therefore any decision made using a myopic policy will not optimize the entire problem. In this sense it would be nice to replace $Q(R_t^x, D_{t+1})$ with an explicit value or approximation at time t. The value function which is discussed in the following section solves just this quandary.

2.9 *Value Function Approximation*

The value function approximation within the aerial refueling model is an estimate of the cost of the receiver refueling and queuing, $Q(R_t^x, D_{t+1})$. The value functions are iteratively created and updated through simulating the cost of refueling receivers with varying levels of tankers. Therefore, the value functions are used in the linear program which incorporates both the explicit first period contributions and the estimation of the second stage contributions (value function approximation)..

The value functions for the aerial refueling problem are used to estimate at time t the downstream value of making decision x_t. This is akin to the decision a New Yorker would make about traveling to a coffee shop. If he standing on a street corner and can walk 1 block west or 1 block east to reach the nearest Starbucks (he is standing on the only street corner in the city without a Starbucks), which location will he choose? Assuming that the explicit costs of moving to either Starbucks location are known to be identical, he is only concerned with the length of the line he will face at each location. Since he has traveled to both locations many times before he has well formed estimates of the which location has the shortest line.

Since the exact total time time (contribution) of moving to either location and waiting in line is unknown at time t, does he just stand on the street corner or make a blind guess about which Starbucks excursion will take the least amount of time? Clearly not, the man walks to the Starbucks he thinks will have the shortest line from his previous experience. The estimate of how long the wait at the two Starbucks will be can be viewed as analogous to the ADP Value Function Approximations!

In the aerial refueling model the same rationale as a man standing on a street corner is used to make the decisions of the tanker movements. When a tanker is sitting at its base and examining the choice of moving to a refueling track, it uses the value of being at the track to guide its decision. Within the linear programming network of Figure 14, the value functions are shown as arcs coming out of the refueling track nodes. Each arc represents the value of having a tanker at that track during the time period. The arc representation is used to convey a more general view of the value function shown in Figure 17, which also shows the value of having additional tankers at a track. As is shown in Figure 17, the more tankers at a track, the less valuable each additional tanker is to the system. The figure is slightly misleading, however, in that it is the slopes of each segment which are important. The slope of each segment represents the value of having having additional tankers at the track. Hence for one tanker the value is the slope of the blue segment (1st segment) while the value of a second tanker at the track is the red segment (2nd segment).

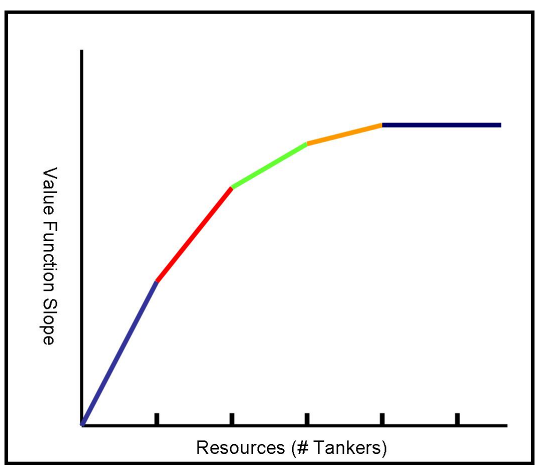

Figure 17: Value Function Approximation

The creation of the value function for the aerial refueling model is again analogous to how the value of a traveling to Starbucks was created by the thirsty coffee drinker. The coffee drinker initially started with no idea of the wait at each location. He essentially started with an empty function (memory) and through repeatedly traveling to each location he was able to create a value for each location. The aerial refueling model also starts with a blank function and no estimation of the value of having tankers at a location, and it uses derivatives from simulation to fill in the function.

In the first iteration there is no known value of having any tankers within the system, and when the linear program is solved no tankers move since only negative cost exists in the system. Since there are no tankers at any of the tracks all receiver missions which enter the system meet a fiery demise. The goal of the aerial refueling model is to reduce the cost of the system, and having receivers crash is an unlikely way to go about optimizing cost in the system. To find the value of having a tanker at a track at time t, the receiver queuing and refueling is re-simulated with the addition of a tanker to the track. The cost associated with receiver queuing and refueling are calculated by the queuing model. The process of adding a tanker to a track and re-simulating the queuing model is repeated for all tracks so that each track has an associated value of having one tanker.

To determine the cost (benefit) of having having the additional tanker the difference between the perturbed and the base simulation within the queuing model is calculated, which is called $\hat{v}_{ta}^{n}$.

$$\hat{v}_{ta}^{n} = C_t(R_t^x + ta, D_t) - C_t(R_t^x, D_t) \qquad (12)$$

Within Equation 12 the value function is identified by the timer period and the iteration of the algorithm.

In the aerial refueling problem once, the value for having an additional tanker at a location is known, $\hat{v}_{ta}^{n}$, it is incorporated as knowledge of the system available in the next iteration. To incorporate the new information into the previously held knowledge an updating formula is used. The updating formula incorporates both the previously known information from prior iterations and the new information learned at the current iteration. The updating formula is:

$$\bar{v}_t^n = (1 - \alpha_n)\bar{v}_t^{n-1} + \alpha_n \hat{v}_t^n \tag{13}$$

Within the value function updating formula the previously incorporated information from prior iterations is identified as $\bar{v}_t^{n-1}$. The $n-1$ identifies that the value function is the smoothed updated from the previous iteration. The incorporation of new information in the value function is guided by the parameter α, which determines the relative weights placed on the previous information and the new information. Alpha is called the stepsize in ADP and the properties of α are further discussed in Section 2.9.2.

The updated value functions from time period t and iteration n, $\bar{v}_t^n$, are then available for use in following iterations to guide the tanker movements. At each iteration and time step the derivatives are calculated around the number of tankers set in the base simulation. When there are tankers at a track during the base simulation, perturbed simulations are run for both one more as well as one less tanker at the track. The derivatives from the perturbations are used to update the value function for having both one more and one fewer tanker. When building a value function, certain states such as having one tanker at a track may be sampled quite frequently while others such as having five tankers may be sampled only once. For the aerial refueling algorithm the value function is only updated at the point where sample realizations occur. More formally:

$$\bar{v}_t^n(r) = \begin{cases} (1 - \alpha_{n-1})\bar{v}_{t,a}^{n-1} + \alpha_{n-1}\hat{v}_{ta}^n & \text{, if } r = R_{ta}^{x,n} \\ \bar{v}^{n-1}(r) & \text{, otherwise} \end{cases} \tag{14}$$

As the algorithm progresses and tankers are assigned to tracks, the importance of having additional tankers at tracks lessens. When the number of tankers at a track reaches a critical mass each additional tanker only decreases the amount of time receivers wait in a queue for refueling. The value function is a concave monotonically decreasing function with respect to increasing resources because of the lessening of the value of each additional tanker. Additionally, since the tankers are indivisible units, the value function is a separable, piecewise linear approximation defined by Equation 15:

$$\bar{V}_t(R_t^x) = \sum_{a \in \mathcal{A}} \bar{V}_{ta}(R_{ta}^x) \tag{15}$$

where $\bar{V}_{ta}(R_{ta}^x)$ is a scalar, piecewise, linear function. The scalar, piecewise function in the aerial refueling model uses the values of the tankers at different track locations and fuel levels to create a value function, an example of which is shown in Figure 17. The value function for the minimization is concave and piecewise linear given the assumptions that for $R_{ta}^x = 0$ the value function $\bar{V}_{ta}(R_{ta}^x) = 0$. Since the value of zero resources is zero that concave function is completely identified by its slopes, which leads to Equation 16.

$$\bar{V}_{ta}^{n-1}(R_{ta}^x) = \left(\sum_{r=1}^{\lfloor R_{ta}^x \rfloor} \hat{v}_{ta}^{n-1}(r) + (R_{ta}^x - \lfloor R \rfloor)\bar{v}_{ta}^{n-1}(\lceil R_{ta}^x \rceil) \right) \tag{16}$$

In Equation 16, $\lfloor R \rfloor$ is the largest integer less than or equal to R, and $\lceil R \rceil$ is the smallest integer greater than or equal to R. The function is therefore completely determined by the set of slopes $(\bar{v}_{ta}^{n-1}(r))$ for all resources from $r = 1, 2, ..., R^{max}$, where R^{max} is the upper bound on the number tankers of a specific type, which for aerial refueling is determined by location and fuel level.

In Figure 18 the idea of the slopes is shown as two different types of tankers value functions overlaid on the same graph. Figure 18 illustrates two different types of tankers at the same location and point in time. In the figure only the fuel levels are different between the tankers such that $X_{fuellevel} > Y_{fuellevel}$. The figure shows both the difference in the value of having additional tankers and also the difference in the value functions of two types of tankers where only the fuel level is varied. When the fuel level is higher each additional tanker has the ability to offload a greater amount of fuel and also each additional tanker has a smaller marginal value. As an example, if there are five receivers at the track with the higher fuel level, the first tanker can refuel three receivers completely. With the addition of a second tanker all five receivers can be refueled, and a third tanker makes it so all receivers can be refueled with zero time spent queuing. For the lower line (tankers with a lower fuel level) the first tanker can only refuel two of the receivers as is the case for the second tanker. Therefore the third tanker refuels the fifth receiver and eliminates any queuing in the system. Hence, the differences in the slopes shown in the overlayed value functions is due to the difference in the marginal value of each additional tanker. The tankers with the lower fuel capacity have a lower value approximation since each of its tankers have less capacity for work than the high fuel level tankers.

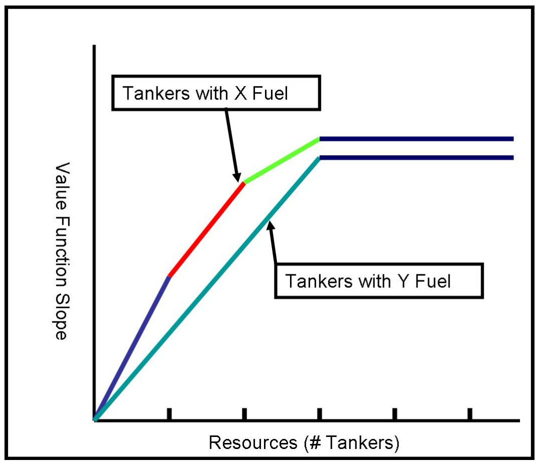

Figure 18: Comparison of Two VFA with Identical Locations and Times but Different Fuel Level Attributes

2.9.1 *Updating and Maintaining the Convexity of the Value Function*

When the derivatives of each resource are calculated, the new value is incorporated into the existing value function for that resource state. As shown previously in Equation 13, a weighted combination of the new value and the previous value of the resource state are used to update the segment of the value function corresponding to that resource level. Since each value function is constructed from a series of approximations about the value of having increasing resources, it is not guaranteed that updating the value function intervals will maintain concavity. Steps must be taken to guarantee that $\bar{v}_{ta}^{n} \geq \bar{v}_{ta}^{n}(r+1)$ for all r when updating a value function approximation interval with a sample value realization $\hat{v}_{ta}^{n}(r) < \bar{v}_{ta}^{n-1}(r+1)$.

The solution to maintaining concavity of the value function is the CAVE algorithm (Concave Adaptive Value Estimation). After the new sample realization information is smoothed into the appropriate interval, the algorithm looks to the left and right intervals to determine if the new function violates concavity restrictions. If concavity is violated then the derivative information is incorporated into the surrounding pieces of the function. The algorithm precedes as follows:

$$\begin{array}{c} \text{if } \bar{V}^n_{t,a}(r) < \bar{V}^n_{t,a}(r+1) \text{than the following smoothing is performed:} \\ \bar{V}^n_{t,a}(r+1) = (1-\alpha_n)\bar{V}^n_{t,a}(r+1) + \alpha_n \hat{v}^n_{t,a}(r) \end{array} \tag{17}$$

$$\begin{array}{c} \text{if } \bar{V}^n_{t,a}(r-1) > \bar{V}^n_{t,a}(r) \text{than the following smoothing is performed:} \\ \bar{V}^n_{t,a}(r-1) = (1-\alpha_n)\bar{V}^n_{t,a}(r-1) + \alpha_n \hat{v}^n_{t,a}(r) \end{array} \tag{18}$$

Equations 17 and 18 are only performed when a concavity violation exists. An example of the updating strategy is shown in Figure 19 for a concavity violation. Without a concavity violation only exponential smoothing occurs (shown in the first three steps of the figure).

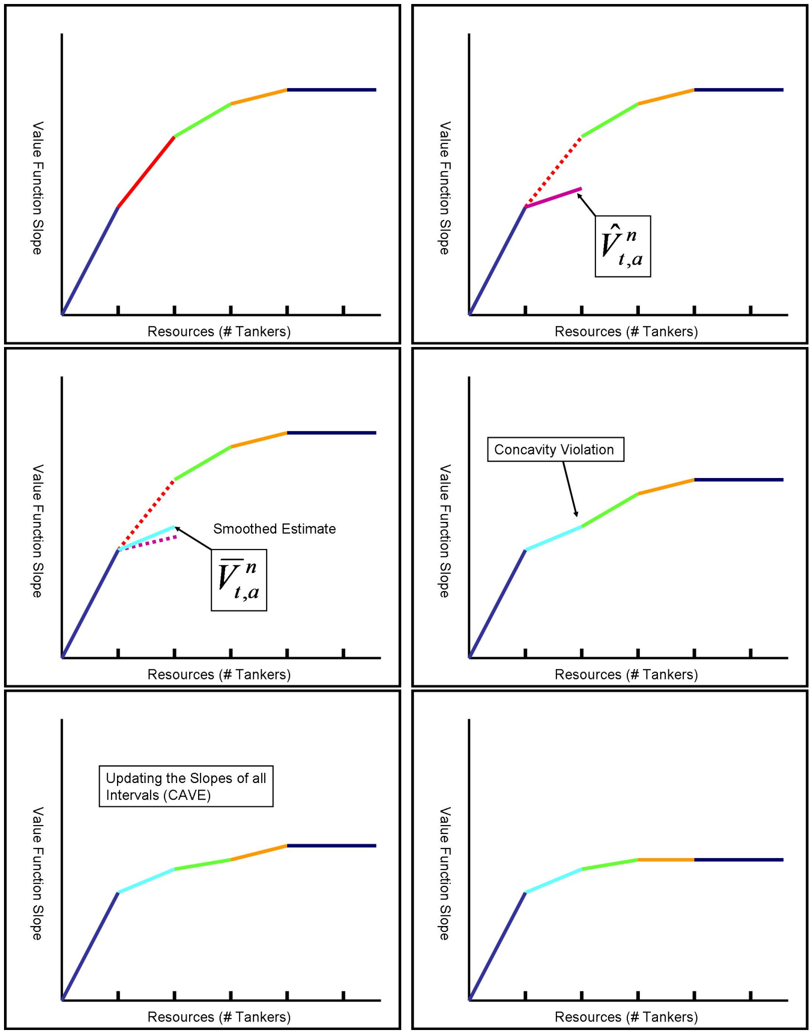

Figure 19: Convex Value Function Adjustment After a $\hat{V}^n_{ta}$

2.9.2 Stepsizes

The variable α plays an important role in updating the value function approximations. The value of α determines the relative weights placed on sample realizations iteration by iteration. The stepsize can impact the convergence of the algorithm since it directly affects value function smoothing. For the aerial refueling model the OSA (Optimal Stepsize Algorithm) stepsize updating algorithm was used due to its ability to incorporate stochastic data and adhere to properties of stepsize algorithms which are provably convergent. The properties of a provably convergent algorithm are:

$$\sum_{n=1}^{\infty} \alpha_n = \infty \tag{19}$$

$$\sum_{n=1}^{\infty} (\alpha_n)^2 < \infty \tag{20}$$

$$\alpha_n \geq 0 \tag{21}$$

A brief explanation will suffice while discussing OSA's use in the current model; however, for a more rigorous discussion the reader is advised to reference Mach Learn (18). The foundation of the OSA is the McClain stepsize size algorithm which is the following:

$$\alpha_n = \begin{cases} \alpha_0 & \text{if n =1} \\ \dfrac{\alpha^{n-1}}{1+\alpha^{n-1}-\bar{\alpha}} & \text{if n} \geq 2 \end{cases} \tag{22}$$

Within the McClain stepsize algorithm the initial stepsize α_0 is set such that in early iterations the stepsize adapts in a similar fashion to the $1/n$ stepsize rule, while in the long run the stepsize approaches a constant stepsize value $\bar{\alpha}$. The OSA algorithm uses the McClain stepsize and modifies it such that it reacts to errors in later prediction with respect to the actual observations. Therefore, while the McClain stepsize naturally decreases throughout the iterations when it is used in the OSA algorithm, it can increase as noise increases and the underlying process shifts and subsequently resumes declining when errors decrease. The behavior of the algorithm allows it to quickly adapt to high levels of noise while also declining to a set stepsize $\bar{\alpha}$.

In a stationary process the stepsizes will decrease toward a fixed value as new data points will provide less and less new knowledge to the system. When the data is highly variable, as with the aerial refueling model in the first iterations, the stepsize will remain high to account for the variability of the information contained in the sample realizations. The variability in the early iterations of the aerial refueling model comes from the high cost associated with mission failures and lengthy queuing. As discussed above, different value functions are created for different fuel levels as well as locations and times. These value functions do not communicate with one another and therefore can be susceptible to large differences in values in reaction to the behavior of other tanker movements.

In an early iteration, if a tanker with a high fuel level and one with a low fuel level are at a track, the tanker with the low fuel level could be given a low value while the high fuel level tanker would have a high value. A later iteration when there is only a single low fuel level tanker at a track without the high fuel level tanker would give the low fuel level tanker a high value for being at that location. By using OSA the difference could be incorporated properly, increasing the value of having the low fuel level tanker, and not mitigated merely because it happens in a later iteration.

2.10 *The Decision Function and the Objective Function*

Having developed the foundations of ADP and their applications for the aerial refueling model, the algorithmic approach for solving the model can be explicitly developed. The contribution function as discussed earlier led to the discussion of using value functions to estimate future contributions. Using the notion of standing at time t and making a decision, which has a known contribution at t and an future unknown contribution at $t' > t$, the decision function is created. Figure 20 shows the linear program which is solved at the beginning of each time step. At time step t, the tankers which are available for movement are the resource nodes. For each resource node all available actions are created and represented in the network as the forward arcs. For these arcs the movements associated with going to a refueling track have value functions. The value functions are represented by arcs, each of which has a value and an upper bound. This is further highlighted in the movements facing a single tanker as shown in Figure 21, where the tanker has five different decision arcs and associated value functions. The decision arc represented without a value function is that of holding a tanker at its base which has no positive value or negative cost associated with the decision.

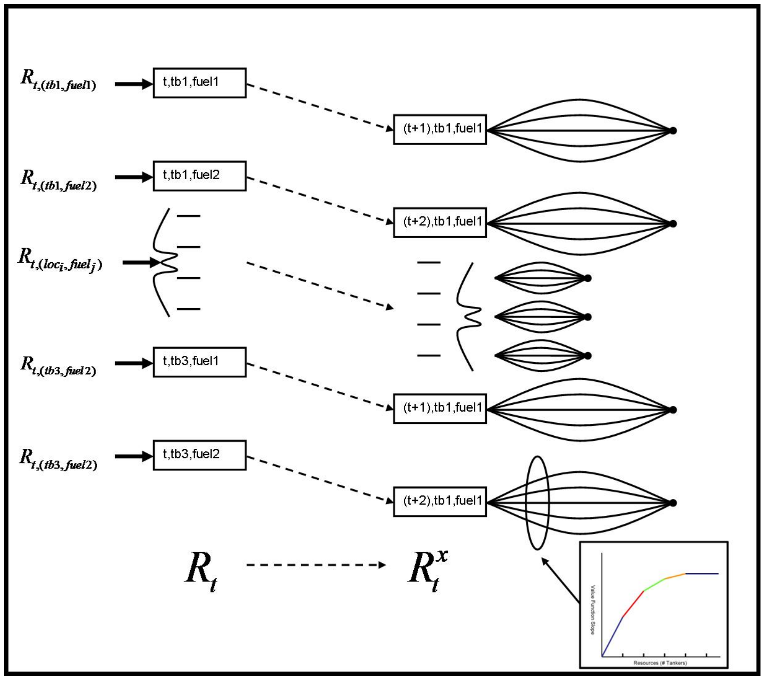

Figure 20: Single Period Linear Programming Formulation with Value Functions

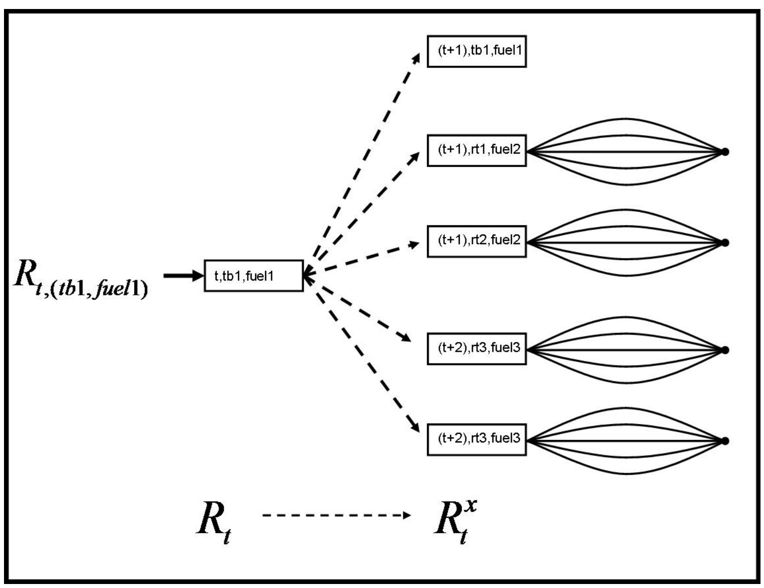

Figure 21: Node Arc Matrix for Single Tanker with Value Functions

As shown in Figures 20 and 21 the tankers have decisions which will take them to both the upcoming time period as well as future time periods. The reason for the different time periods is the amount of travel time required for a tanker's movement from its current location to the various refueling track locations. Additionally, this means that the contribution function in Equation 11 which was assumed to take the immediate contribution and the next period's contribution, is in actuality more complicated than looking one period into the future. A more representative contribution function for a movement is:

$$C_t(R_t, x_t) = \sum_{a \in \mathcal{A}, d' \in \mathcal{D}} c_{tad} x_{tad} + Q(R^x_{t'}, D^x_{t'}) \tag{23}$$

In the above equation $t' > t$ and t' also represents the last time period before another tanker decision has been made on tankers moved initial at time t. More explicitly, since value functions represent the future value of having a tanker at a location at time t, a tanker "sees" the queuing value previously computed from a similar tanker at an earlier iteration (similar fuel level and location). While future contributions are explicitly calculated at a future time period and applied to that period, they are used in a previous time period to make decisions.

The decision and delayed contribution is very similar to that of filling out a W-2 and filing taxes. At the beginning of a year an individual can choose to withhold money for taxes throughout the year or defer any withholding and pay the full tax bill at the end of the year. While withholding payments or the lump payment happen in the future at time period $t = 0$, a decision must be made which is binding throughout the year. If the lump payment is chosen then throughout the year the tax payments which have been deferred can be invested in T-Bills. At the end of the year, for the lump payment option, the contribution to wealth is the difference between the tax payment and the growth of the invested deferred tax payments which have been in T-Bills. The contribution to wealth which occurs at time $t = 12$ is a direct result of a decision which occurred 12 periods before. Therefore, it is not unreasonable to say that the contribution from the decision at $t = 0$ is the immediate contribution and the end contribution even though it isn't realized for 12 periods since no other decisions have occurred in the interim. While the bank does not record any increased wealth until the end of the year, it can be assumed by the decision maker to have happened much earlier. This is how the aerial refueling model works, in that the cost of queuing is recorded in the total cost of the simulation when it actually occurs, but the cost of queuing for solving the decision function is associated with the decision in a previous time period.

For the aerial refueling model to solve the optimal decision, the best policy is found by searching over the group of policies, $X_t^\pi(S_t)$, and solving the equation:

$$\max_{\pi \in \Pi} \mathbb{E} \sum_{t=0}^{T} \gamma^t C_t(S_t, x_t) \tag{24}$$

The aerial refueling model uses a simple myopic policy where the contributions from each individual point in time are maximized. The optimization problem for the aerial refueling model is represented by:

$$X_t^\pi(S_t) = \arg\max_{x_t \in \mathcal{X}_t} \sum_{a \in \mathcal{A}, d \in \mathcal{D}} C_t(a_t, d_t). \tag{25}$$

Solving the optimization problem in Equation 24 for the aerial refueling model means solving a series of myopic linear programs. The myopic policy is determined through the linear program which maximizes the linear programs objective function. Within the objective function the cost of fuel associated with moving a tanker/holding a tanker at a refueling track are negative values. The value function arcs in the linear program are calculated as positive values. When the derivatives of having tankers at a track are smoothed into the value functions the decrease in cost from having additional tankers is either positive or zero. Therefore, the model looks at the cost of moving a tanker to a track versus the benefit of having that tanker at the proposed track and solves Function 25 through optimizing the objective function accordingly.

2.11 *The Algorithm*

To solve the aerial refueling problem a forward pass algorithm shown in Figure 22 is used. The forward pass algorithm uses value functions from the previous iteration to make its decisions. At the end of an iteration the value functions are updated accordingly and available for use in the following iteration.

Step 0: Initialization:

Step 0a. Initialize $\bar{V}_t^0, \quad t \in \mathcal{T}$.

Step 0b. Set $n = 1$.

Step 0c. Initialize R_0^x (The set of all tankers in the system).

Step 1: Choose a sample realization, ω^n. For $t = 1, 2, \ldots, T$, (ω is the deterministic list of receiver missions in the aerial refueling simulations) do:

Step 2a: Create the linear program from the available tankers and associated value function approximations:

Step 2b: Solve the optimization problem:

$$\max_{x_t \in X_t^n} [C_t(R_t^n, x_t) + \bar{V}_t^{n-1}(R^{M,x}(R_t^n, x_t))]$$

Step 2c: Simulate the receiver refueling and queuing to find $\hat{v}_t^n(R_t^x)$

Step 2b: Increment $R_t^x \pm \epsilon$, at all tracks.

Step 2d: Re simulate the queues with the $\pm\ \epsilon$ to find the derivatives which are $\hat{v}_t^n(R_t^x(\pm\epsilon))$

Step 2e: If $t > 0$ Update the appropriate value function using:

$$\bar{v}^n(r) \quad = \quad \begin{cases} (1 - \alpha_{n-1})\bar{v}_{t-1,a}^{n-1} + \alpha_{n-1}\hat{v}_{ta}^n & \text{if } r = R^n \\ \bar{v}^{n-1}(r) & \text{otherwise} \end{cases}$$

Step 2f: Update the States:

$S_{t+1}^n = S^{M,W}(S_t^{x,n}, D_{t+1}, W_t)$

Step 3. Increment n. If $n \leq N$ go to **step 1**.

Step 4: Return the value functions, $\{\bar{V}_t^n, t = 1, \ldots, T, a \in \mathcal{A}\}$.

Figure 22: An approximate dynamic programming algorithm to solve the aerial refueling problem.

3 *Receivers Falling Out of the Sky!!(Does the Model Work?)*

Having the general framework for the aerial refueling model established, the actual implementation of the model into a working simulator that will provide reliable, efficient results becomes the focus of the rest of the paper. What defines whether the model is optimizing and providing usable solutions? The initial focus is guaranteeing that the model can quickly and reliably reduce mission failures to zero. Mission failures occur if a receiver is not assigned to a tanker when it enters the model. For the model to be usable and provide a feasible solution, mission failures must be eliminated. In many ADP models, satisfying all demands is not necessary in determining the validity of the model; however, the aerial refueling model must consistently eliminate mission failures to be of any value to operators of the model.

Once the model has been shown to consistently reduce mission failures to zero then the ability of the model to optimize costs is the next goal of the system. The model is designed to reduce the total cost accrued through tanker and receiver movements and refueling. The aerial refueling model is expected to have high mission failure and queuing cost in initial iterations; however, through the use of value functions the tanker movements should be optimized and lower the cost of a simulation throughout the iterations. The costs associated with various aspects of the model such as tanker fuel, receiver fuel, and queuing should be optimized in concert throughout the optimization without any one cost dominating to the detriment of another cost.

The third goal of the model is to produce reliable results which make sense and are usable by mission planners. Example of this goal include: reducing total tanker usage to a minimum and consistent level when given an excess amount of tankers in the system, reducing individual receivers' queuing times to acceptable levels, and refueling receivers at logical locations. The usability of the model for the Air Force requires that these goals are met, and while the model may be correct in all technical dimensions, without results which mirror those expected by planners it may be considered useless.

To achieve all of the goals of the model, the inputs and structure of the model were required to closely mimic the real world with regards to actions and decisions. The following sections detail the model-specific attributes of the aerial refueling simulator which help it mirror the real world.

3.1 *Modeling With Realism*

The aerial refueling model implements a series of constraints and changeable parameters to make the actions of the tankers and receivers more realistic. To model the tankers, the fuel levels of tankers are accurately updated throughout the simulations. Additionally, decisions are guided through policies which limit the actions of tankers as fuel levels deplete. Such a measure includes limiting tanker movements at an epoch to returning to base immediately if the tanker does not have enough fuel to stay on station for another time interval and return home with a safe margin of fuel. Another constraint put on the tankers guarantees that tankers will reject refueling any receivers that will deplete their fuel to a level which will not allow the tanker to return home with an adequate level of fuel. This constraint has the dual role of guaranteeing that tankers return home and also that receivers are not assigned to tankers that would be forced to return home while the receivers are still waiting in a queue.

A tunable parameter for the tankers is the turn around time associated with a tanker returning home to base. Tankers that return to base after refueling are unusable for at least four hours, which mirrors refueling and crew changes as well as guaranteeing that one tanker is not expected to be airborne 24 hours straight. Another added benefit of a long turn around time is that the model is forced to efficiently allocate and move tankers. When the holding time of a tanker returning to base is combined with the traveling time associated with returning to base the tanker leaving its track is unavailable to return to a track for upwards of seven hours. Therefore, anytime that the model sends a tanker to base it is unavailable to refuel receivers at a track for upwards of ten hours. By limiting the missions each tanker can refuel in a day, the stress on the system was increased and conservatively reflected how often a tanker can be used daily.

The last major constraint to the system is the refueling time for the receivers. The refueling time for receivers is an endogenous constraint of the system. The refueling time for aircraft is set such that there is a margin of error for when the plane can be refueled; however, with fighter and attack planes such as the F-18 and F-15 the limited excess fuel carried on board relative to fuel burn rate demands that they refuel at or close to the specified time. While the goal of the model is to eliminate queuing, the current Air Force model has a built-in 15 minute window that allows tankers and receivers to wait before attaching and refueling. The leeway allowed in the current Air Force model is incorporated into the aerial refueling model by stipulating that planes incur no penalty for refueling under 15 minutes after their scheduled time and incur penalties for delays past 15 minutes. By allowing for minimal delays the model closely mirrors the actualities of refueling while not penalizing the inherent stochastic nature of refueling times. The penalty as well as the time limit are both exogenous variables and thus can be adjusted to suit the user's desires; however, the current implemented values balance receiver failure and fueling delay cost.

After implementing all of the major required constraints into the system, the model optimized the aerial refueling problems and did so in a manner that compared favorably with the current Air Force planning model. In the next section the tunable inputs and outputs of the model are discussed to guarantee the reader is familiar with the world of aerial refueling and the inner workings of the CASTLELAB Aerial Refueling Model.

3.2 *The Results*

To accurately gauge the success of the aerial refueling model, the current Air Force model provided by Jim Donovan from AFOSR was used as a baseline. Throughout the early testing, Mr. Donovan's Excel-based model was used as guidance on the number and location of tankers required to adequately serve all the receiver missions. Once the number of tankers required to solve the receiver mission profile in Mr. Donovan's model was ascertained, the current model results were shown to approach and improve upon those results. The results from runs of the AFOSR model are in Tables 2 and 3. As discussed earlier, the Excel model's optimization capability is limited since it pairs of tankers to receivers in a strictly myopic fashion. Another constraint on the Excel-based model is that the receiver refueling tracks are endogenous to the system. Therefore, the AFOSR model is limited because it optimizes only the tanker movements while taking the receiver movements as fixed inputs. The model developed in CASTLELAB therefore cannot mimic the results of the AFOSR model. A limited comparison between the aerial refueling model and the AFOSR model using the SDS showed the aerial refueling model requiring 16 tankers while the AFOSR model required 20 tankers. Since a direct comparison of the models was not possible this baseline test which showed that the aerial refueling model produced similar results to the AFOSR model is used to illustrate the general validity of the ADP approach.

Simulation	Tanker Base	Given KC-10A	Tankers Used
1	BASE 1	20	20
2	BASE 2	20	20
3	BASE 3	20	18
4	BASE 4	20	18

Table 2: Tankers Used by AFOSR Model for Varying Tanker Inputs

Simulation	10 Tankers KC-10A	10 Tankers KC-10A	Used Base A	Used Base B
1	BASE 1	BASE 3	8	10
2	BASE 2	BASE 4	8	10

Table 3: Tankers Used by AFOSR Model for Varying Tanker Inputs

After the validity of the aerial refueling model was established in comparison to the AFOSR model, a series of tests were run on the aerial refueling model to establish the characteristics and strengths of the model. The results are framed in the context of producing a usable model for the Air Force, and therefore, some of the tests were established to test the usability of the model while other tests were performed to determine the robustness of the model.

3.3 *The Model Inputs*

To test the aerial refueling model, two distinct data sets were used which provided insight into different aspects of the model. The first data set used is a small data set (SDS) consisting of 4 tanker bases, 4 receiver bases, 4 tracks, and 58 missions. The second data set (LDS) is a much more complex data set with 5 tanker bases, 14 receiver bases, 19 tracks, and 117 missions. Both data sets cover missions over a 24 hour horizon. The major difference in the complexity of each system involves the differences in the number of tracks in the sets. The number of tankers and receivers in the system provide limited computational complexity since only distances traveled and fuel burns must be calculated. However, the VFA are measured at tracks, and by increasing the number of tracks there is a direct increase in the intricacy of the problem as each track must account for a variety of value functions at each time step to account for different tankers. Therefore, the LDS is a much richer data set than the SDS, and the results of the LDS can be considered more applicable to the real world except in a few examples.

To test the LDS, a number of inputs were used to create a base case scenario as listed in Table 4.

Variable	Iterations	Tankers	Rcvr Penalty	Fuel Ratio	Movement Penalty
Base Set	100	25	10,000	2.18	0.6

Table 4: Base Data Set Inputs- LDS

- Iterations-The number of iterations the simulator was run.
- Tankers-The tankers within the system (all tankers are equally distributed throughout tanker bases during test runs).

- Receiver Penalty-The model-specific penalty for a mission failure. Receiver missions which are not refueled during an iteration are defined as failures. The Receiver Penalty is also used in the computation of the cost of a receiver fueling delay. Receiver fueling delay is defined as the time a receiver sits in a queue.

- Fuel Ratio-Importance of tanker fuel usage relative to receiver fuel usage. The base case is with tanker fuel burn rate set at 14,400 pounds/hr and receiver fuel burn set at 6,600 pounds/hr, which are values taken from Air Force refueling manuals. The model therefore initially values a tanker in the air costing 2.18 more per hour than a receiver.

- Movement Penalty-A receiver mission's distance traveled is broken into two legs - base to track - track to target. The second leg of the receiver mission costs more than the first due to the receiver wanting more fuel in the combat zone on its way to its target and therefore can be penalized.

3.4 *The Model Outputs*

The outputs measured in the simulations focused on a variety of metrics which are important to the Air Force planners, as well as statistics which show how well the model is optimizing. The model outputs for the Air Force focus upon the fuel burned within the system, the fueling delay encountered by the receivers (queuing time), the number of tankers used in the system over the complete time horizon, and the distance traveled by the receivers.

The fuel burned is separated into two categories, the fuel burned by the receivers and the fuel burned by the tankers in the system. In addition to the fuel used in the system, the total cost of the system includes the cost of mission failures as well as total fuel burned. Figure 23 is an illustration of the fuel burned throughout the iterations for the base LDS simulation. When measuring the system, the solution is not considered stable if mission failures occur after the initial learning iterations; therefore, the total cost of the system is only measured when the system is stable.

When the system is not considered stable, the results will note the instability and the outputs should be taken with caution.

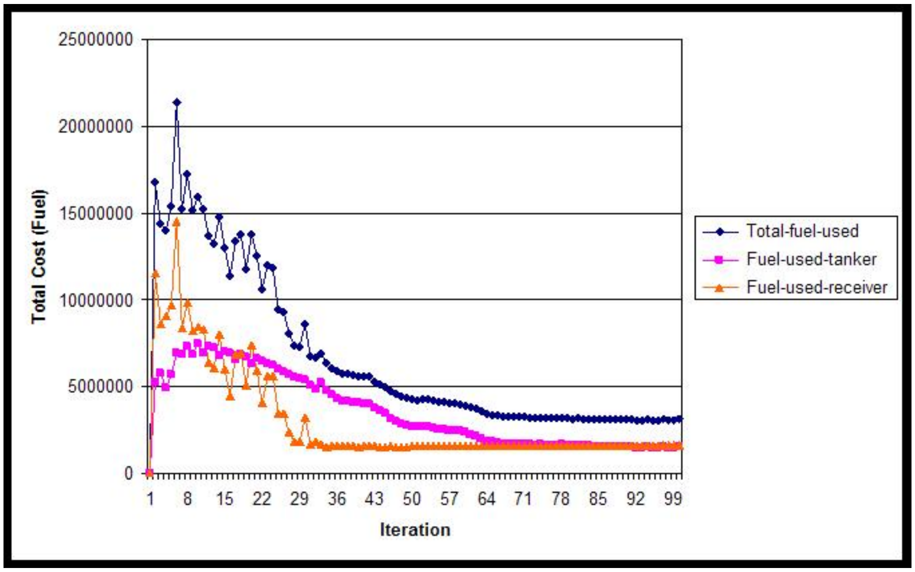

Figure 23: Total Fuel Used in Pounds for the Base LDS Simulation

The fueling delay is measured with several metrics. The first measure is that of total fueling delay within the system. This measure is important since it is an indirect measure of how flexible the system is to added receiver missions and imprecise fueling times. When the total fueling delay is low, the measure shows that there is little overlap of assigning receivers to identical tankers at the same time which produces queuing. Therefore, introducing instability (real world frictions) to a system with low total queuing would have a lower impact on the system than simulations that have a large fueling delay. The other measure of fueling delay focuses on the maximum delay encountered by any single receiver in the system. When the fueling delay encountered by a single receiver is large, $delay > \mathcal{X} minutes$, a penalty is assessed to the system as the receivers do not have a large excess fuel capacity. The model is set to minimize fueling delay for each receiver and an acceptable delay is defined as lasting under 15 minutes. An example of the optimization of total fueling delay in minutes per iteration is shown in Figure 24.

The total tankers required in the system throughout the time horizon and the

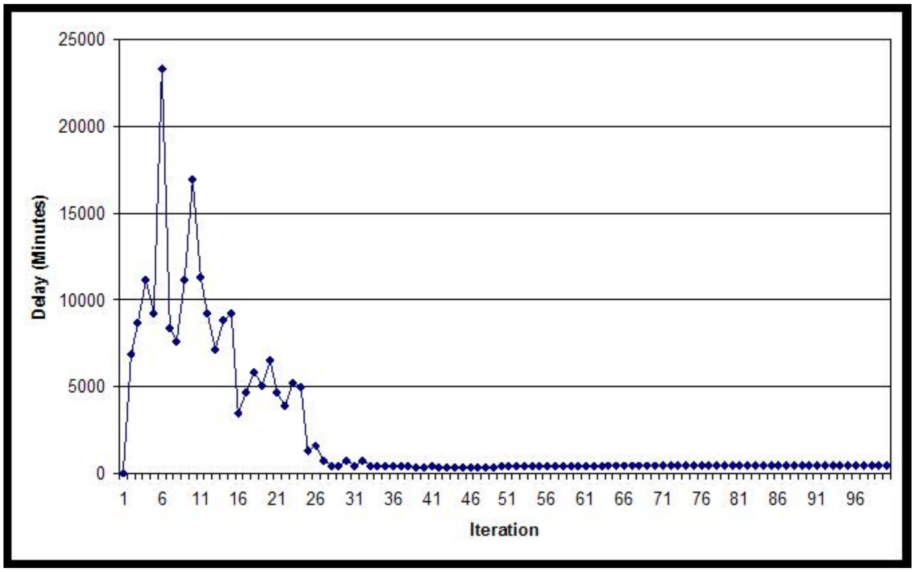

Figure 24: Total Fueling Delay for the Base LDS Simulation

efficiency of tanker usage in the system are also measured. The measure of tankers required in the system is important since it shows the minimum amount of tankers required in each iteration to produce the given results. Throughout the iterations the expectation is that the tankers required by the model decrease to a stable value. Figure 25 illustrates how the base LDS uses all the available tankers (25) for the first 60 iteration before "learning" that it can produce a better solution with fewer tankers. The aerial refueling model is set up such that if two identical tankers are sitting at a base and one of the tankers has previously been used (flown to a refueling track and then back to base) then the previously used tanker will be reused in the model. The tie breaking rule guarantees that the aerial refueling model uses the minimum number of tankers required and does not unnecessarily fly previously unused tankers.

The second measure of how tankers are used is the tanker usage efficiency which focuses on how well the model optimizes the tanker movements in the system. When a tanker moves from its base to a track, it is moving due to the perceived value of the move which is from the VFA. However, given that the VFAs are not exact predictors of the future they can cause moves which have no value. As the algorithm

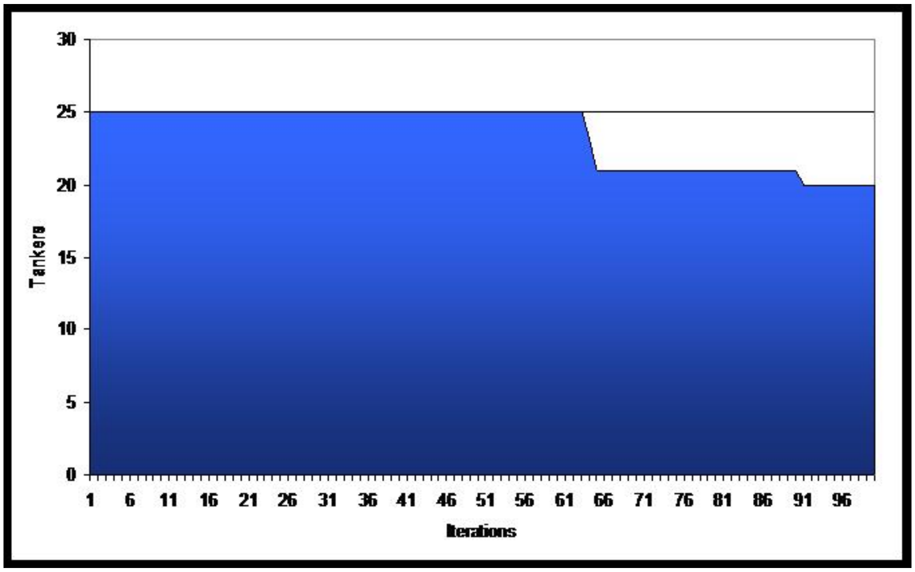

Figure 25: Total Tankers Used Per Iteration for the Base LDS Simulation

progresses, unnecessary moves by the tankers should decrease as the value functions become more refined. The measure of the average number of tankers at a track during an iteration shows how many tankers the system has moved from base to a track or are held at a track due to a perceived value of having tankers at the track. The measure of the average number of tankers unused at a track shows the number of tankers which were sent to a track and subsequently were not used for refueling any receivers. The average unused tankers in the system are expected to steadily decline during the iterations as value functions become more accurate and send the appropriate number of tankers to the correct refueling tracks. Additionally, as the average of unused tankers decreases, the average number of tankers at a track will decrease since tankers are used more efficiently. As shown in Figure 26, in early iterations there are excess tankers both used and unused at tracks, but during the later iterations the used tankers reach a steady value and the unused tankers approach zero as tanker movements are optimized.

The final measure of the system comes through the total objective function cost associated with an iteration. The total objective function cost is a measure of how well the model is optimizing the total cost of the system in the linear program. Through

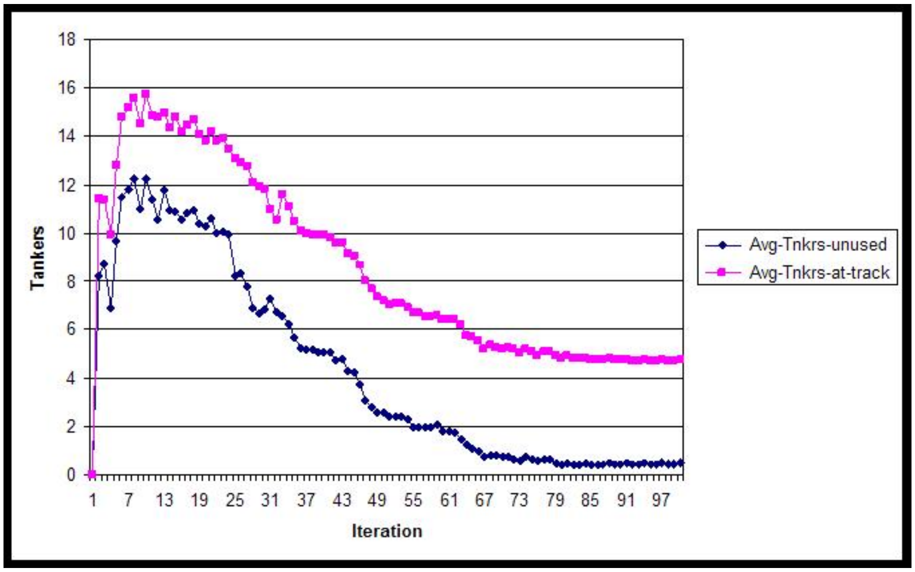

Figure 26: Average Tankers Used Per Time Step in an Iteration for the Base LDS Simulation

the iterations, as the value function approximations improve and tanker assignments become more precise, the objective function decreases. The objective function is a composite of the contribution from moving a tanker to a track and the value function approximation associated with that movement. In Figure 27 the initial high objective value is due to exploration and imprecise value function approximations; however, as the iterations progress the objective function settles into a stable region which is around the optimal objective value. In our simulations the optimal objective value is not computable as the state space is too large. As a proxy, the percentage change in the objective function between iterations is computed and used to measure of stability of the model. As shown in Figure 27, the objective function is very stable over the last 50 iterations.

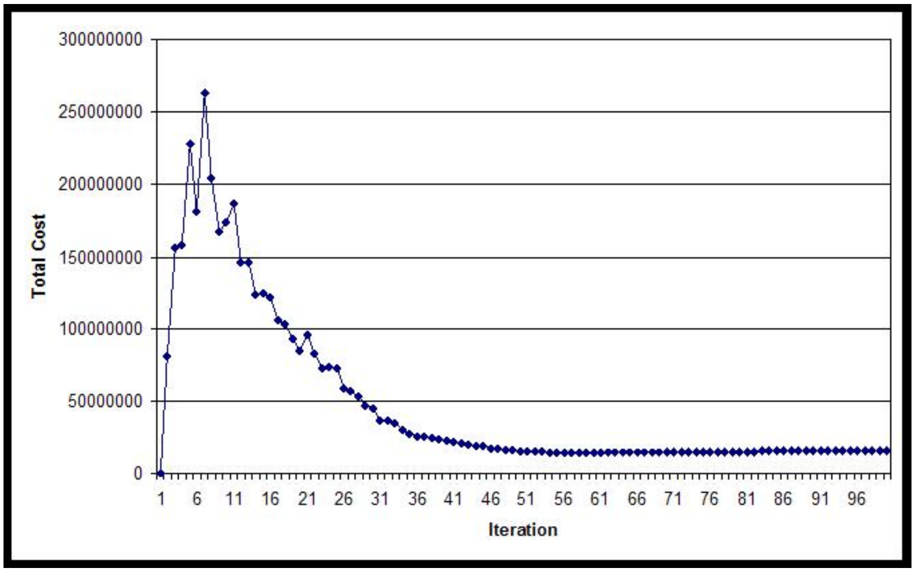

Figure 27: Total Objective Function Cost for the Base LDS Simulation

3.5 *How Quickly Does the Model Work?*

When testing the model, the speed of the convergence of the solution is an important metric. As stated above, the absolute convergence to a known optimal value is not possible. Rather, the relative changes in the objective function are used to determine the stability of the solutions. The stability of a solution is important over a long horizon in ADP due to the common occurrence of relative convergence. Relative convergence occurs when an algorithm is run over a short horizon until the solution appears to reach an optimal solution, but it has in fact reached a sub optimal solution which would become obvious with more iterations. When examining Figure 28 it appears that the solution is stable around 40 iterations.

However, when than simulation is extended to 100 iterations, as shown in Figures 29 and 30, the solution and equilibrium of the solution changes quite a bit. The first figure (29) shows the total cost across all of the simulations and the second figure (30) illustrates the total cost change between the 40th and the 100th iterations. The second figure clearly illustrates that the solution improves and converges on a solution that was not apparent when the simulation was only run for 40 iterations. Therefore,

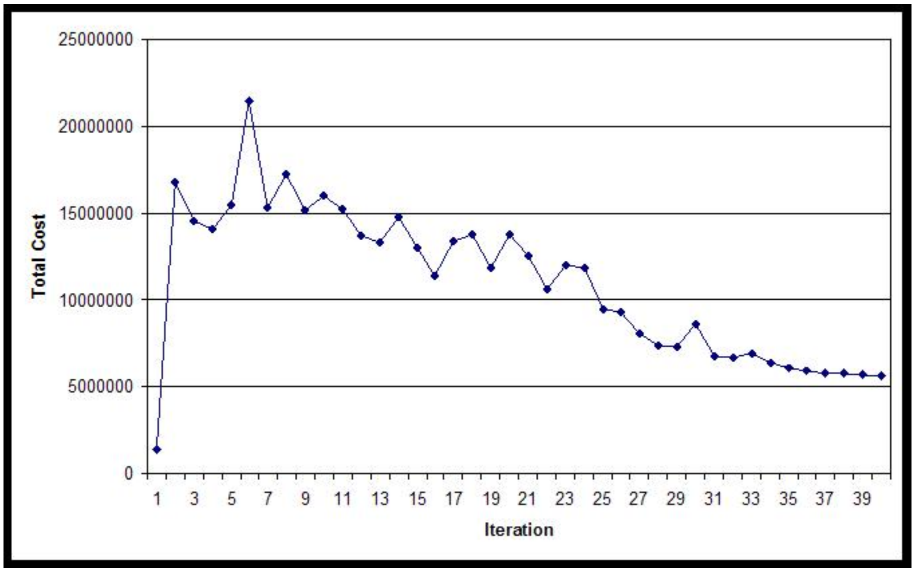

Figure 28: Total Cost - Apparent Convergence over First 40 Iterations

it is important to find out how quickly the solutions converge to a stable solution which persists over an extended horizon.

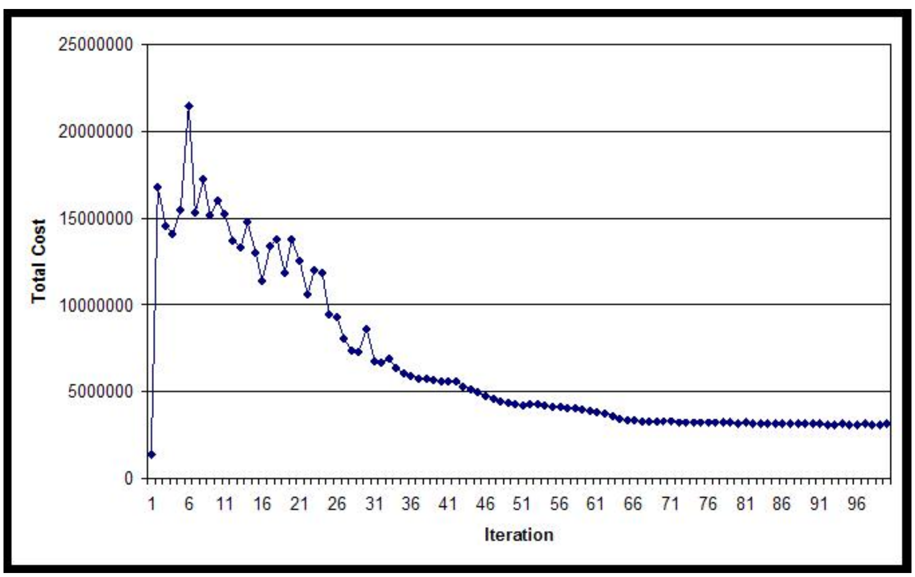

Figure 29: Total Cost - Apparent Convergence over First 100 Iterations

Using the following inputs for the large and small data sets (Figures 3.5 and 3.5), the optimal simulation length concerning the trade off between the stability of the solution and the memory and time required to run the simulations was established.

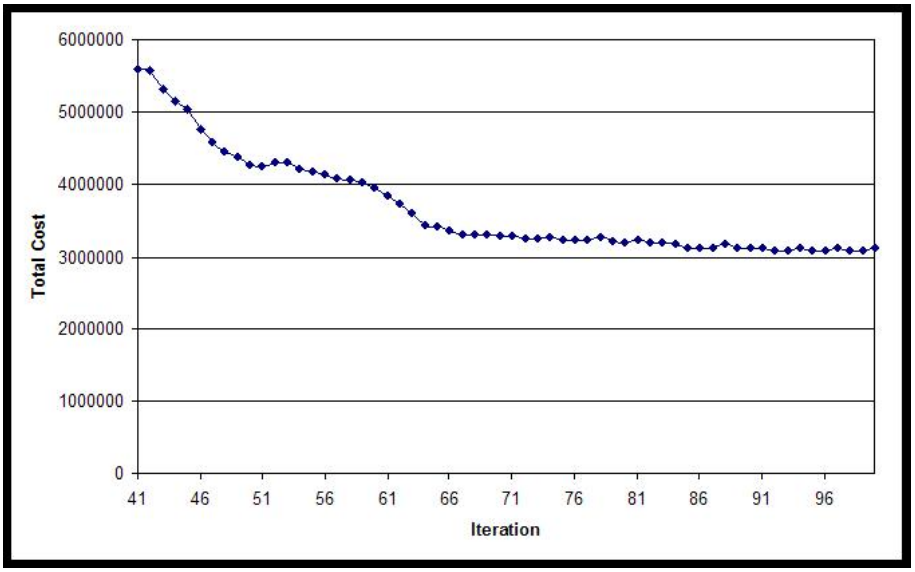

Figure 30: Total Cost - Apparent Convergence from Iteration 41 to 100

The differences between the LDS and the SDS in terms of iterations required are due to the difference in the measured state space of the two data sets. As noted earlier, the LDS and SDS states are measured at discrete intervals with regard to the location of tankers, receivers, and the various states of each of the resources and demands. While the SDS and LDS have similar amounts of tankers, there is a large difference in the number of locations between the two sets. The LDS has more than four times the tracks contained in the SDS data set (19 vs 4) and thus the LDS-measured state space and value functions are more than four times as great as the SDS. Therefore, the LDS requires more iterations to reach a stable solution than the SDS.

In the aerial refueling model, one state of the world at each time step of an iteration can be explored. Therefore, in the first iteration the value of having one tanker at each track is calculated through creating derivatives and updating the associated value functions. The second iteration uses the value function approximation from the first iteration to determine where to place the tankers in the second iteration. The third iteration uses the information gained in the previous two iterations to move tankers in the system and so forth. When the number of tankers in the system is less than the number of value functions, there is a limit to the state space which can be explored in an iteration, and subsequently a limit to the number of value functions which can be updated. As tankers attempt to update the various value functions by exploring the state space, the algorithm is said to be in an exploration phase. With a large state space (LDS) the exploration phase of the ADP algorithm is much longer than in a more compact state space (SDS). As shown in the outputs and graphs of the base LDS (Table 3.5 and Figure 31) and SDS (Table 3.5 and Figure 32) data sets there is a great difference between the rate of convergence between the two sets, which is expected due to the difference in the states spaces explored.

Variable	Iterations	Tankers	Rcvr Penalty	Fuel Ratio	Movement Penalty
Set 1	20	25	10,000	2.18	0.6
Set 2	50	25	10,000	2.18	0.6
Set 3	100	25	10,000	2.18	0.6
Set 4	200	25	10,000	2.18	0.6

Table 5: Large Data Set Inputs - Varying Simulation Length

	RcvrFuel	TankerFuel	Delay	MaxDelay	TnkrUsed	Unused	Used
Set 1	3444314	6023105	1346	14.33	25	8.23	13.08
Set 2	1582003	2691280	437	14.33	25	2.58	7.17
Set 3	1595082	1525753	486	11.33	20	0.50	4.75
Set 4	1583113	1577220	535	11.33	19	0.17	4.17

Table 6: Large Data Set Outputs - Varying Simulation Length

Variable	Iterations	Tankers	Rcvr Penalty	Fuel Ratio	Movement Penalty
Set 1	20	20	10,000	2.18	0.6
Set 2	50	20	10,000	2.18	0.6
Set 3	100	20	10,000	2.18	0.6

Table 7: Small Data Set Inputs - Varying Simulation Length

For the LDS after examining the tradeoff between the rate of change of the total cost and the time required the standard simulation run was set at 100 iterations. The SDS converges much more quickly than the LDS and the standard simulation length was set at 50 iterations.

	RcvrFuel	TankerFuel	Delay	MaxDelay	TnkrUsed	Unused	Used
Set 1	3712778	2315654	434	32	16	.25	4
Set 2	3712778	2315654	434	32	16	.25	4
Set 3	3712778	2315654	434	32	16	.25	4

Table 8: Small Data Set Outputs - Varying Simulation Length

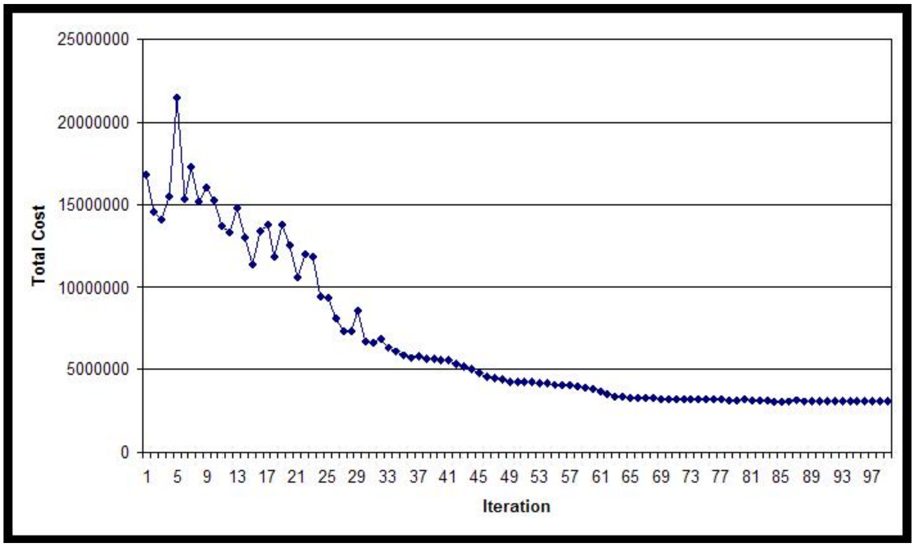

Figure 31: Total Cost for LDS

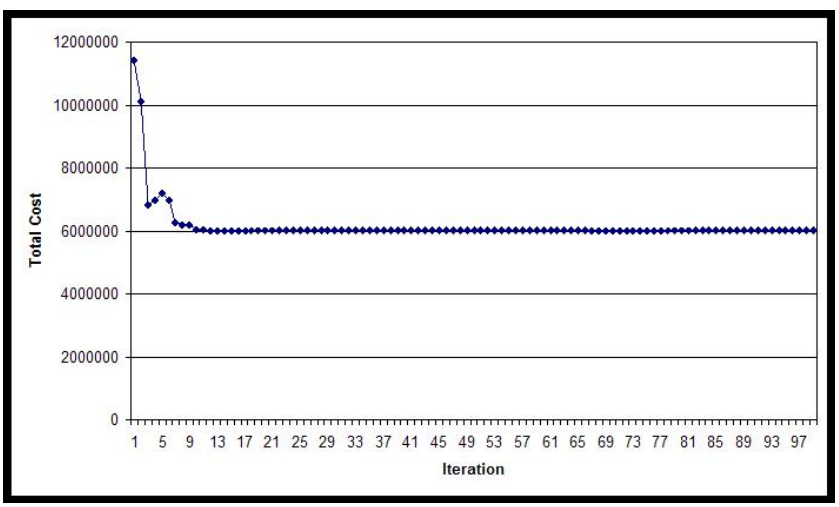

Figure 32: Total Cost for SDS

3.5.1 *The Importance of Quickly Obtaining Stable Solutions for the US Air Force*

The US Air Force is concerned with planning missions in a time efficient manner which can be updated daily if not more frequently. Using data from past engagements of the United States military, the daily receiver missions during Operations Enduring Freedom and Iraqi Freedom can reach over 1,000 in a day, as shown in Table 1 in Section 1.4. The daily receiver mission rate is therefore eight times larger than the LDS. A model which requires too many iterations, and therefore computing time, would be of limited use to the Air Force planners as they must set forth a schedule daily and be able to deal with uncertainty and change the model as necessary throughout the day. The amount of iterations required to reach a stable solution in the aerial refueling algorithm is more responsive to refueling tracks and tankers in the system than receivers at any given point. Therefore, a model which has a similar structure and size with regards to available refueling tracks and tankers could be solved in a similar number of iterations. The time required to run one iteration of the LDS is 25 seconds,which involves invoking a remote linear programming solver (CPLEX) while using an older desktop machine running at 1.5 GHz. As most machines which would run this software would be faster than the test machine, there is an expectation that the scalability of this algorithm to the full data set is not a limiting issue.

Additionally, as will be discussed in much greater detail in Section 4.4, the algorithm can be set up to run in a "warm start" state which uses previously trained value functions. Therefore, for the LDS a single run of 100 iterations can be used to train value functions, and the trained value functions can be used to run a similar data set and produce solid results in five to ten iterations.

3.6 *The Value of Tankers in the System*

Approaching the aerial refueling problem with the ADP algorithm required the examination of the solution quality for a series of inputs. The most important input to be able to change while maintaining solution quality is the number of tankers in the system. The algorithm should be able to use various numbers of tankers and produce solutions which are similar given the changing tanker inputs.

Differing levels of tankers are able to sample the state space more or less completely during each iteration due to the availability of tanker resources. However, it is expected over a long horizon of iterations that all levels of tankers will explore the state space and create similar value function approximations. The creation of similar value functions for varying levels of tanker will confirm the validity of the model. It is important that the varying levels of tankers produce similar results so that that model is not dependent upon the skill of the operator in determining the number of tankers required by the system prior to a simulation.

In the Air Force there are established guidelines for assigning tankers to receiver missions; however, the approach of the aerial refueling model takes a much different tack. A strength of the model would be that it can optimize the system regardless of the number of tankers input by an inexperienced user. A naive approach to assigning tankers to the system by an inexperienced mission planner does not focus upon mission efficiency, but rather is concerned solely with guaranteeing receiver mission completion. Using a naive approach, the optimal level of tankers is unknown and the level of tankers assigned to the system will likely be much greater than the required level of tankers. A model that can produce similar solutions both when an model operator assigns close to an optimal level of tankers as well as when they assign a great excess of tankers would show the ability of the aerial refueling model to optimize. Additionally, the flexibility of the aerial refueling model would provide a great level of usability to operational planners.

Testing both the LDS and SDS with varying levels of tankers, the conclusions detailed in Sections 3.6.1 and 3.6.2 highlight the algorithm's ability to optimize with varying levels of tankers.

3.6.1 *Optimizing With Tankers Assigned To All Tanker-Bases*

To test the ability of the model to react to varying levels of tankers, multiple simulations were run in which differing numbers of tankers were placed in the system and simulated (all tankers were distributed equally amongst the tanker bases). Additionally, the system was set up such that at each tanker base location there was a virtually unlimited number of tankers, (25). As shown in the base LDS run (Table 3.5), the model required 20 tankers to successfully refuel all receiver missions; therefore, each tanker base location alone could successfully refuel all of the receiver missions. The test of the model was to check whether the algorithm would be able to optimize over a larger state space of tankers and come up with a solution which used a similar number of tankers as the base LDS simulation (20). Additionally, it was expected that the other output metrics in Table 4 would be similar in scale. As the results from Tables 10 and 12 show, as the number of tankers introduced to the system increased the fuel cost and tanker usage statistics were lowered for both the LDS and SDS when compared to the base simulations.

Variable	Iterations	Tankers	Rcvr Penalty	Fuel Ratio	Movement Penalty
Set 1	100	15	10,000	2.18	0.6
Set 2	100	25	10,000	2.18	0.6
Set 3	100	50	10,000	2.18	0.6
Set 4	100	100	10,000	2.18	0.6

Table 9: Large Data Set Inputs when varying Tankers

	RcvrFuel	TankerFuel	Delay	MaxDelay	TnkrUsed	Unused	Used
Set 1	3761080	4031937	1974	627	15	5.12	8.38
Set 2	1595082	1525753	486	11.33	20	0.5	4.75
Set 3	1583113	788610	535	11.33	20	.17	4.17
Set 4	1537087	897554	529	11.33	19	.25	4.33

Table 10: Large Data Set Outputs when varying Tankers *note Set 1 is unstable with mission failures after 100 iterations

The dramatic decrease in the fuel consumption for both the LDS and SDS between Sets One and Two (Table 10) and Sets Three and Four (Table 12) is due to the model

Variable	Iterations	Tankers	Rcvr Penalty	Fuel Ratio	Movement Penalty
Set 1	50	16	10,000	2.18	0.6
Set 2	50	20	10,000	2.18	0.6
Set 3	50	32	10,000	2.18	0.6
Set 4	50	50	10,000	2.18	0.6

Table 11: Small Data Set Inputs when varying Tankers

	RcvrFuel	TankerFuel	Delay	MaxDelay	TnkrUsed	Unused	Used
Set 1	3721521	2427170	434	36	16	.25	4
Set 2	3712778	2315654	434	36	16	.25	4
Set 3	3730493	1812812	434	36	18	.25	4
Set 4	3730493	1702683	434	36	18	.25	4

Table 12: Small Data Set Outputs when varying Tankers

optimizing movements of tankers from closer tanker base locations. Since there are more tankers at tanker bases that are close to highly used refueling tracks, the tankers from the close bases are used and tankers from bases farther away are not required. The use of more “local” tankers as the tankers at each base are increased explains the large decrease in the total tanker fuel consumption. Ignoring LDS Set 1 due to its instability from a lack of tankers, it is clear that for the LDS and SDS simulation runs the receiver fuel burn remains relatively unchanged among all the sets. The stability of the receiver fuel burn shows that the assignment of receivers to refueling tracks is consistent once a critical mass of tankers are in the system. This is consistent with the approach taken to estimate the value function approximations and receiver assignment rules.

An interesting and yet counterintuitive result of the simulations is that the receiver fuel consumption decreases to a stable value much more quickly in the sets with many tankers than in sets with fewer tankers, as shown in Figure 33 for the LDS. Intuitively, the data sets with fewer tankers allow less freedom of operation for the receivers, as they can refuel at fewer refueling tracks, and thus the receivers' fuel burn rate would be expected to converge at a faster rate. However intuition is misleading with respect to the aerial refueling algorithm.

The sets with greater levels of tankers can more quickly explore a larger section of the state space in fewer iterations. During the initial "learning" iterations the data sets with more tankers are able to send tankers to more of the available tracks than the sets with few tankers. Since tankers are assigned to more tracks, the value function approximations associated with the "best" tracks are updated more frequently in early iterations. This is due to receivers having a simple decision function of moving to the track which has a tanker and produces the shortest distance from base-track-target. When there are limited tankers in the system some of the "best" tracks will not be sampled during the initial exploration phase. With a limited number of tankers in the system there is a constant pull between exploration and exploitation of the state space. Even with a limited set of resources, eventually the tankers can sample a large portion of the state space and reach a solution which is similar to the data sets with greater levels of tankers. Figure 33 illustrates this point clearly since all three data sets from the LDS converge on similar values, but their rate of convergence varies greatly.

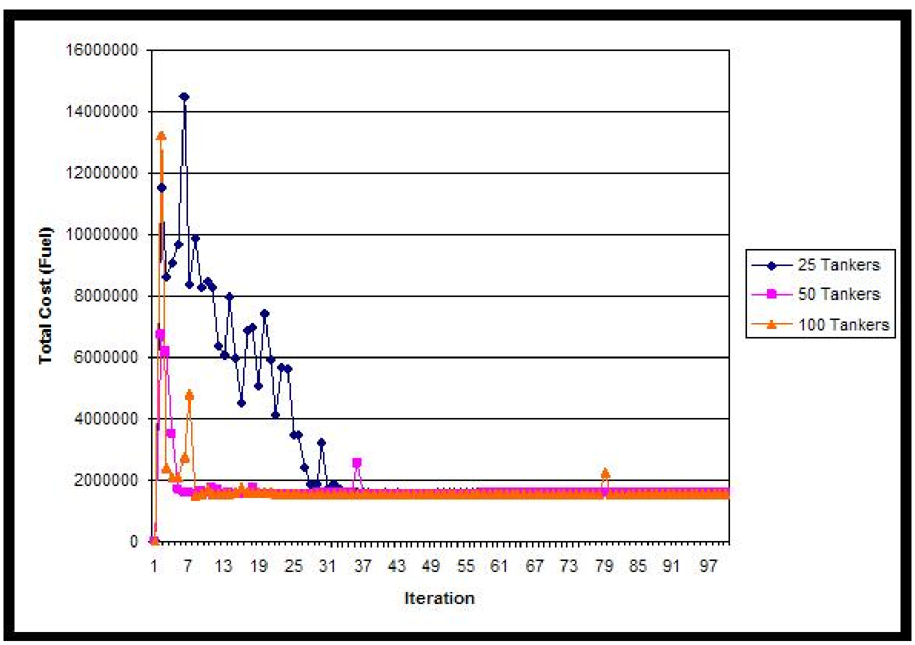

Figure 33: Receiver Fuel Consumption Comparison with Varying Levels of Tankers for the LDS

As discussed above, the total tanker fuel burn rate varies greatly since the required tankers fly from more favorable tanker bases; however, as shown in Figure 34 there is more to the solution than simply the distance tankers must fly. The results and conclusions are similar between the LDS and the SDS, but the LDS more clearly illustrates the conclusions due to its larger state space. Figure 34 shows the differential tanker fuel consumption totals between the LDS data sets. For the different sets there are two distinct phases which are the initial 10 iterations and then the subsequent 90 iterations. Within the first ten iterations it is expected that Set 3 and Set 4 would send out more tankers than Set 2, and therefore their fuel burn rates would be higher than Set 2. The graph shows that in the initial ten iterations it is the case that the sets with more tankers have greater fuel consumption; however, after ten iterations the set with fewer tankers is burning much more fuel than the other sets. After the first 15 iterations, Set 3 and Set 4 are approaching their optimal fuel burn rates while Set 2 is still in its exploratory phase. As discussed above Set 2 has fewer tankers and thus it takes more iterations than Sets 3 or 4 to explore the state space sufficiently and determine its optimal decisions. Therefore, it takes Set 2 longer to reach its equilibrium, and at equilibrium there is the added complication of having to send tankers from more distant locations so it has a higher optimal tanker fuel burn cost.

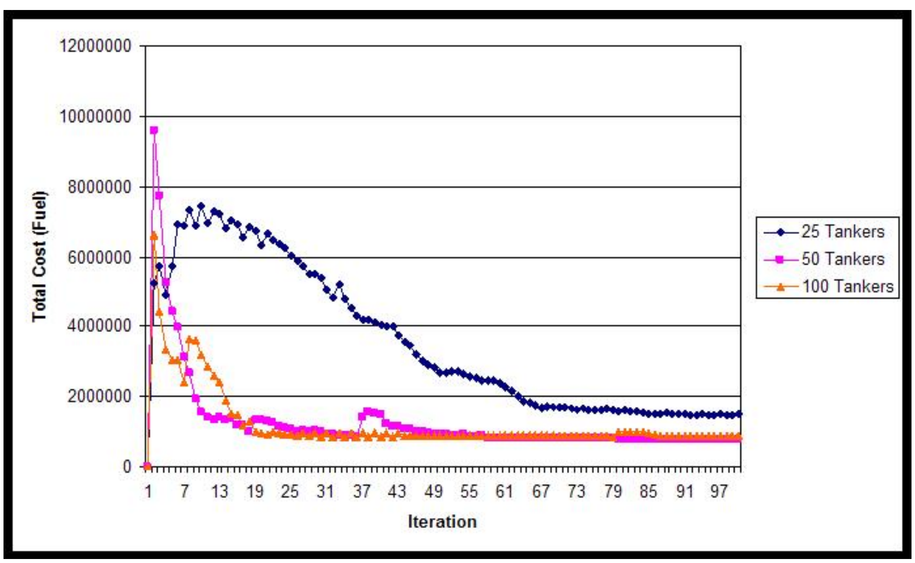

Figure 34: Tanker Fuel Consumption Comparison with Varying Levels of Tankers for the LDS

The sets with large tanker fleets send most of their tankers from a small subset of the available tanker bases. With the larger fleets at each tanker base the model can move all tankers the shortest possible distance without having to pull tankers from the second choice (longer distance tanker base). Set 2 must move tankers from multiple bases to fill a demand at a single track and when this is accounted for the rate of convergence is slowed. Additionally, since the tankers are pulled from bases which are farther away than the optimal tanker base, more fuel is burned. Therefore, the large difference in the tanker fuel consumption after 100 iterations is a function of the distances flown by the available tankers and to a smaller extent, the slower rate of convergence.

3.6.2 *Optimizing With All Tankers at a Single Tanker-Base*

The model has been shown to pick the most desirable tankers when there are tankers at multiple locations, but another important attribute of the model is optimizing over a fleet of tankers at a single location. The previous section showed that a tanker fleet given an excess of tankers will choose the most desirable tankers based on location and availability, but how well does the model optimize when tankers are only at a single location?

To test the ability of the model to optimize over a single location, two locations within the LDS were chosen and given 100 tankers for separate simulations. The two tanker base locations were chosen for their relative closeness to the refueling tracks used in the base LDS simulation. Location A is closer to the aerial refueling tracks in the base LDS simulation than Location B. It is expected that Location A will more quickly send out tankers due to the decreased movement cost of tankers to refueling track when compared to Location B. However, as the simulations progress the movements of tankers from both Location A and Location B, as well as the total cost, are expected to be similar.

As shown in Figure 35, Location A optimizes much more quickly then Location

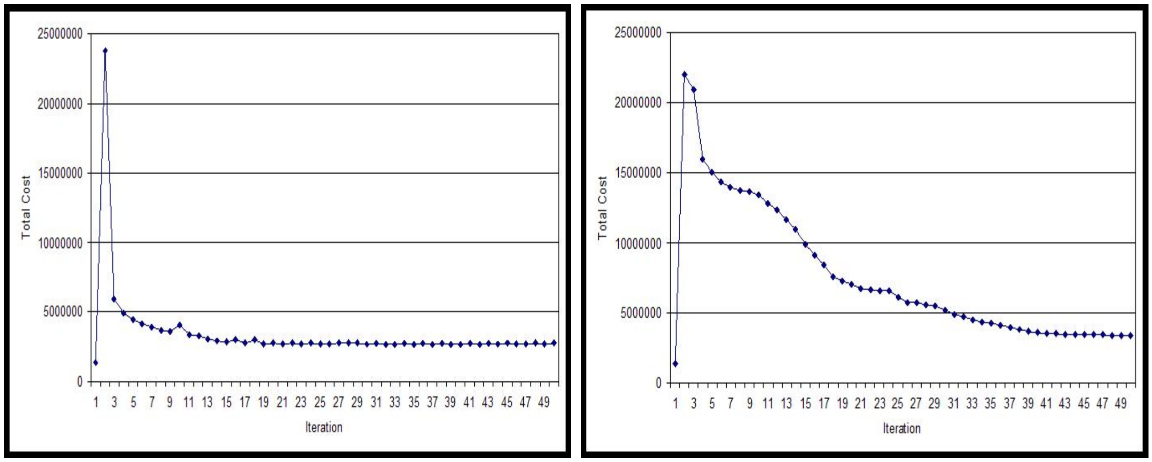

Figure 35: Total Cost Per Iteration For Location A (Left) and Location B (Right) for the LDS using 100 Tankers at a Single Tanker Base

B. Since the linear program at the heart of the model is constructed of both the value function approximations and tanker movement cost this is an expected result. In the early iterations, the tankers at Location A have a very low cost associated with moving to refueling tracks, while those from Location B have a much higher cost for moving. The lower threshold for moving tankers causes more tankers to move to the refueling tracks in early iterations and thus an optimal solution is found more quickly.

Location B has a higher cost threshold for moving tankers to tracks and thus in the first iterations it moves fewer tankers. By moving fewer tankers to tracks in the first four iterations, the values built in the VFAs for having one or two tankers at a track is very high as many receivers fail. Figure 36 shows that after the fourth iteration the value of moving tankers to tracks has become high enough to move a majority of the tankers from Location B to refueling tracks. Since in the early iterations the value functions at all refueling tracks consistently showed receiver mission failures, the model must then recompute value functions at all refueling tracks as tankers move to the refueling tracks in later iterations. The smoothing associated with this calibration of the value functions slows the convergence for the simulation of Location B. However, as the simulation progresses both locations use a similar number of tankers. Both simulations also have similar total cost, but the cost of sending the tankers to tracks from a more distant tanker base is reflected in the slightly higher cost of Location B.

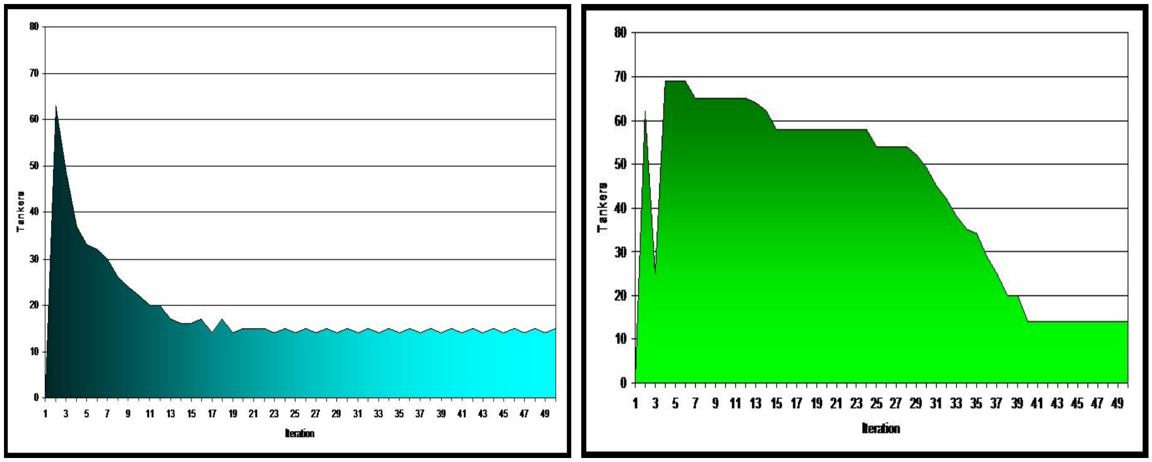

Figure 36: Tanker Usage Per Iteration for Location A (Left) and Location B (Right) for the LDS using 100 Tankers at a Single Tanker Base

The results of the aerial refueling model when a single tanker base location is used mirror those expected in real life. When a lower cost is associated with a move it requires much less value to make the move positive. Therefore, the quick convergence of Location A to a stable value is expected. For a longer move, as with Location B, it takes a higher value to make a move a positive choice. The model works in this manner for Location B as it requires the value functions to build high values before moving tankers. Also, the model is responsive to the many value functions which exist within the system. In the early iterations for the Location B simulation, positive values are built at many refueling tracks due to continuing mission failures. In the other simulation, as there no mission failures in early iterations due to optimal tanker placement, the value functions at tracks without tankers are updated with a value of zero for having one tanker. Therefore, for the Location A simulation, the linear program does not send tankers to unused tracks after the initial iterations since there is not a positive value associated with the moves. Conversely, in the Location B simulation, the artificially high value function approximations from the early iterations must be corrected through the system "learning" the correct placement of tankers and values associated with those locations. As the system learns the correct locations the values associated with having tankers at unused locations decreases to a low enough level that tankers are not longer sent to those locations.

The aerial refueling model can consistently optimize from a single location as well as from multiple locations. Additionally, increasing numbers of tankers in the system are handled by the model and can dramatically decrease the iterations required to reach optimality. The consistent results which occur when varying the number of tankers in the system show that the value function approximations are insensitive to tanker inputs. Therefore, the stability of the value functions highlight the usability of the model for mission planers since the model's results are not dependent upon any operator skill or finesse.

3.7 *The Value of Fuel*

The purpose of this model is to minimize the fuel cost associated with refueling receiver missions for a given set tankers. Therefore, it is important that the fuel burn characteristics of both the tankers and the receivers accurately reflect the rates of planes in the Air Force inventory. Throughout this research a constant, specific fuel burn rate for both tankers and receivers in the system was used. While there are added complexities to the fuel burn rates of planes such as differential rates between take off, cruise, and refueling, the complexities were ignored for the sake of concise, applicable results. In the model, the tankers burned fuel at the rate of 14,400 lbs/hr and receivers at 6,600 lbs/hr, which were values derived from "AFPAM 10-1403, AIR MOBILITY PLANNING FACTORS" used by the US Air Force when making gross calculations of aerial refueling requirements.

Built into the aerial refueling model is the implicit assumption that when making decisions for tanker movements and receiver movements, moving a tanker is 2.18 times more expensive than moving a receiver. The fuel ratio, fr, is the burn rate of the tanker divided by the fuel burn rate of the receivers.

$$fr = \frac{burn_{tanker} lb/hr}{burn_{receiver} lb/hr}. \tag{26}$$

Since tankers are assumed to burn fuel at a rate which is 2.18 times greater than the receivers, the model will likely choose shorter movements for the tankers and move the receivers greater distances. This solution appears to be out of line with the dynamics of the real problem, where receivers have far less fuel than tankers and therefore each pound of their fuel is more valuable. By changing the cost associated with burning tanker fuel in the model, the results will provide insight into where receivers would refuel if tanker movements through the system are essentially cost free.

Given that in the model a higher value is placed on tanker fuel than receiver fuel, it was determined that the cost of tanker fuel would be dropped such that it would be less costly to fly an hour in a tanker than a receiver. The lower fuel burn rate is only incorporated in the explicit movement cost of the tankers and not in calculating actual fuel burned, which updates the attribute vector of the tanker. By only changing the cost of a tanker movement, the dynamics of how long a tanker can be in the sky or the amount of receivers a tanker can refuel are not changed, but rather only the cost associated with moving a tanker in the linear programming formulation. In Tables 13 and 14, Fuel Ratio is the cost associated with the fuel burn rates between the tankers and the receivers. When the fuel ratio is set at 0.1; the model assumes the receivers burn fuel at a rate which is ten times costlier than the tankers.

Variable	Iterations	Tankers	Rcvr Penalty	Fuel Ratio	Movement Penalty
Set 1	100	25	10,000	0.10	0.6
Set 2	100	25	10,000	1.0	0.6
Set 3	100	25	10,000	2.18	0.6

Table 13: Large Data Set Inputs with Changing Fuel Ratios

Variable	Iterations	Tankers	Rcvr Penalty	Fuel Ratio	Movement Penalty
Set 1	50	20	10,000	0.10	0.6
Set 2	50	20	10,000	1.0	0.6
Set 3	50	20	10,000	2.18	0.6

Table 14: Small Data Set Inputs with Changing Fuel Ratios

The data sets reacted differently to varying the fuel burn rates and therefore the conclusions and limitations of this approach are discussed in two parts.

3.7.1 *LDS Results*

Lowering the tanker fuel burn rates did not provide improved solutions for the LDS. Within the model there is a commingling of the tanker and receiver fuel burn cost as well as the receiver mission failure cost, which complicates the expected results of the model when changing the tanker fuel burn cost variable. Table 15 shows that when the fuel burn rate for the tankers is lowered (Set 1 has the lowest cost), the receivers and tankers actually burn more fuel and the solution is unstable due to continuing mission failures. In addition to the increased fuel consumption, the model optimizes much slower and continues with a large number of unused tankers after 100 iterations.

	RcvrFuel	TankerFuel	Delay	MaxDelay	TnkrUsed	Unused	Used
Set 1	5,206,932	5,124,072	3308	718	25	9.44	13.19
Set 2	1,985,168	2,951,200	646	140	25	4.17	9.58
Set 3	1,595,082	1,525,753	486	11	20	0.5	4.75

Table 15: Large Data Set Outputs with Changing Fuel Ratios

The explanation for the failure of an improved receiver solution with a lower tanker fuel burn cost is rooted in the fuel burn rates of the receivers themselves. The tanker movement decisions occur in the linear program. In the LP the cost of moving a tanker is compared with the value associated with having a tanker a track. The value of moving the tanker to a track is determined from the value function approximations. In the LDS base configuration, all of the input variables work in concert and reliably decide when tanker should move to a track. However, when the tanker fuel cost is reduced greatly for the LDS the decisions are much less reliable for two reasons.

The first reason the results suffer stems from the decreased threshold for sending a tanker to a track. In the aerial refueling model, queuing under 15 minutes is not penalized and therefore the only savings from sending an additional tanker to a track with a receiver queue is the savings gained from reducing the queuing fuel burn cost to zero. Considering a queue of ten minutes and the standard receiver fuel burn rate of 6,600 lb/hr, the savings of an additional tanker which eliminates the queue is 1,100 pounds of fuel. In the base LDS simulation a tanker would not move to save the system 1,100 pounds of fuel unless the distance was less than five minutes away, since in five minutes the tanker would burn 1,100 pounds of fuel. Therefore, in the base model the receivers would enter a queue and be served by the original tanker. When the tanker fuel burn rate cost is dramatically decreased to 660 lb/hr, the dynamics of the model change considerably. With the lowered fuel burn rate the tanker can travel up to 100 minutes to eliminate queuing and will have burned the same amount of fuel as the queueing it eliminates. With the lowered threshold for sending tankers to tracks to reduce queuing the model sends out most of the available tankers in early time steps of an iteration. The movement of the tankers in the early time steps results in less tanker availability in the later time steps as the tankers are sitting at their bases refueling and receiving maintenance. The lack of tankers in later time steps accounts for the dramatic increases in queuing that occurs in later time steps of an iteration.

The second reason that the results do not improve when the tanker fuel cost is lowered is that there exists many more tanker movement decisions which have similar fuel burn cost. This is important because normally there are distinct choices when comparing distances due to fuel burn rates. When the tanker fuel burn cost is lowered it changes the scale of the comparison between fuel burn rates and mission failure cost. Therefore, through this lack of scale more tankers enter the system than should for a certain level of receivers. When the tanker fuel burn cost is at a more reasonable 6,600 lb/hr ($fr = 1.0$), the problem is not as dramatic as at the lower cost of 660 lb/hr but it still exists. Examining the results when the Fuel Ratio is 1.0 the solution is heading in the correct direction; however it is taking dramatically longer to reach an optimal solution then the standard fuel burn ratio of 2.18. In the SDS results section 3.7.2 the outputs are more in line with the base outputs; however, the results are more indicative of a smaller state space which will be discussed below.

3.7.2 *SDS Results*

The SDS suffered from the commingling of variables as in the LDS; however, this is mitigated due to the SDS having only four tracks and the long distances associated with reaching those tracks. Within the SDS the distances traveled to tracks by tankers are much greater than those in the LDS. The average tanker base to track distance for the SDS is 1,054 miles while in the LDS it is only 606 miles. The increase in distance makes it far less attractive to move tankers to save queuing time in the SDS than in the LDS. In the SDS, a plane must queue for nearly double the time of the LDS before it appears attractive to move a tanker and save the queuing time. Additionally, the increase in the fuel required to travel home decreases the amount of time tankers in the SDS are able to stay on a track, regardless of the tanker fuel burn cost. Since the distances are greater, tankers are forced to return home instead of staying at a track. In the SDS the problems associated with the LDS are diminished due to the unique structure of the data set; however, even with this data set the results don't show a marked decrease in the total fuel burned by the receivers, as shown in Table 16.

	RcvrFuel	TankerFuel	Delay	MaxDelay	TnkrUsed	Unused	Used
Set 1	2,506,515	1,314,679	116	12	19	0.25	4.38
Set 2	2,554,547	1,909,531	116	12	20	0.12	4.25
Set 3	2,487,073	2,131,740	116	12	16	0.12	4.25

Table 16: Small Data Set Outputs with Changing Fuel Ratios

The results from the LDS and the SDS show that changing the cost of the tanker fuel to an artificial level does not affect the total receiver fuel burn cost dramatically, but can introduce problems within the model. Changing the tanker fuel burn cost to lower levels in the LDS caused tanker behavior which had negative affects on both receiver and tanker fuel burn cost. The SDS does not suffer from the shortcomings of the LDS due to its structure, but it was shown that changing the tanker fuel cost did not noticeably decrease the total receiver fuel burn cost of the system. Additionally, the LDS is a much richer data set and more instructive of the results which would be expected of other large data sets. Therefore, while it superficially appears that reducing the tanker fuel burn cost would produce a better receiver solution, it is shown to have little upside but a large possible downside, and it is not recommended that attempts at changing the behavior of tanker movements through changing fuel burn cost to artificial levels are instituted.

3.8 *Moving Planes on Target with Maximal Fuel Loads*

The previous section examined the differences in the total receiver fuel burned when the cost of tanker fuel is lowered. The previous approach was not very instructive for a variety of modeling reasons, and its use would have been of limited value in real world situations. A major limitation to artificially changing the tanker fuel burn cost is that in the real world supply officers want to minimize fuel burn by both entities. In this section another approach at influencing receiver behavior without artificially altering fuel costs is shown.

When a receiver mission takes off from its base the first leg in its mission is reaching the refueling track and linking with a tanker. After finishing the first leg of the trip the receiver moves from refueling track to the target. Within the mission, the fuel level of the receiver has much greater value during the second leg than the first. There are several reasons for valuing fuel to a greater extent in the second leg of the mission, which involve the ability of the receiver to move at high speed if necessary (which has a higher fuel burn rate), the face that more fuel allows the receiver to patrol for targets of opportunity, and a greater initial fuel load ensures that a receiver will have adequate levels of fuel to exit the combat zone. Since the fuel level is more important in the second leg than the first, it is reasonable to assume that a solution which refuels receivers closer to their intended targets is one goal of mission planning.

The aerial refueling model incorporates a scaling factor on the second leg of a receiver mission which can be tuned to make flight profiles with shorter track to target distances preferred to profiles with longer track to target distances. Below is shown the exact type of behavior the scaling factor will produce and the simulation results of implementing the scaling factor.

In Figure 37 the distance profiles of a plane flying to a target via Track 1 and Track 2 are illustrated. For this example both the tanker and the receiver are launched from the same base and must travel to either Track 1 or Track 2 to refuel the receiver. When comparing the fuel burn of the receiver between traveling to Track 1 and then on to its target, or to Track 2 and then on to its target, the differences appear negligible with Track 2 holding a slight advantage. However, since the model optimizes over the total fuel burned in the system, the fuel burned by the tanker is also considered when picking the optimal track.

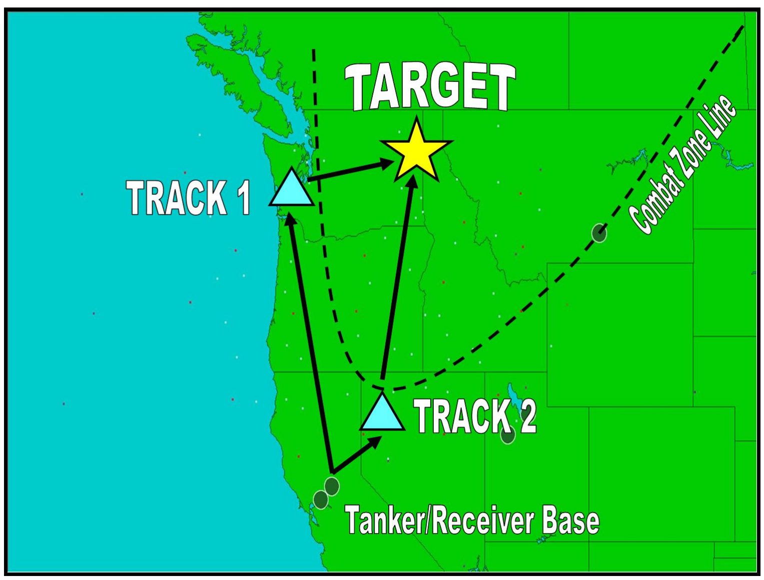

Figure 37: Track Distance Movement Example for Two Tracks

Flying a tanker to Track 1 involves a much longer tanker round trip flight than flying to Track 2 and therefore the fuel cost is much greater. The minimization of fuel cost for this brief example is simply calculated as the combined fuel burn of the tanker and receiver and Track 2 is the obvious preferred choice. While Track 2 is the best choice for minimizing fuel cost in this example, the solution ignores any outside influences for which Track 1 might be preferred to Track 2 in spite of the increased fuel cost. In certain situations it is not unreasonable to assume that Track 1 is preferred to Track 2 since the receiver will enter the combat zone with far more fuel, but how can the aerial refueling model ever chose Track 1 without hard coding the model with data set specific rules?

The answer is the previously mentioned approach of separating the receiver mission profile into two distinct parts. In the aerial refueling model the receiver's flight distance is broken into two components: the flight from base to the track and the flight from the track to the target. By placing a penalty factor, x, on the second leg of the trip when the receiver decisions are made, it can be assigned to the track with a tanker which is closest to its target. While this appears to be a brute force method, it actually is quite subtle in its execution since tanker movements are directed solely through movement cost and value functions. The value functions which are used to decide where to move tankers can be influenced through the method of splitting the receiver movement into two parts during the early iterations. During the early iterations which are purely exploratory, the model places tankers at all the available track locations subject to tanker constraints. In these early iterations, influencing where the receivers travel also influences how the value functions are built at locations. Equations 27 through 30 govern the total cost of the system and are shown below:

$$C_{tnkr} = 2 * D_i \tag{27}$$

$$C_{rcvr} = D_i + (1 + x) * D_{i,target} \tag{28}$$

$$C_{total} = C_{tnkr} + C_{rcvr} \tag{29}$$

$$i \in \mathcal{I} = \text{Set of all track locations} \tag{30}$$

In Figures 38 - 40 an example problem is shown to illustrate the influence that changing the value of the penalty factor, x, can have on the movements of receivers and tankers in the system. Within the system there are two tankers and a single receiver. In iteration A (Figure 38) there are no tankers at either track but a derivative is calculated at each track for having a tanker and the value functions are updated.

With the updated value functions in the second iteration (Figure 39), tankers fly

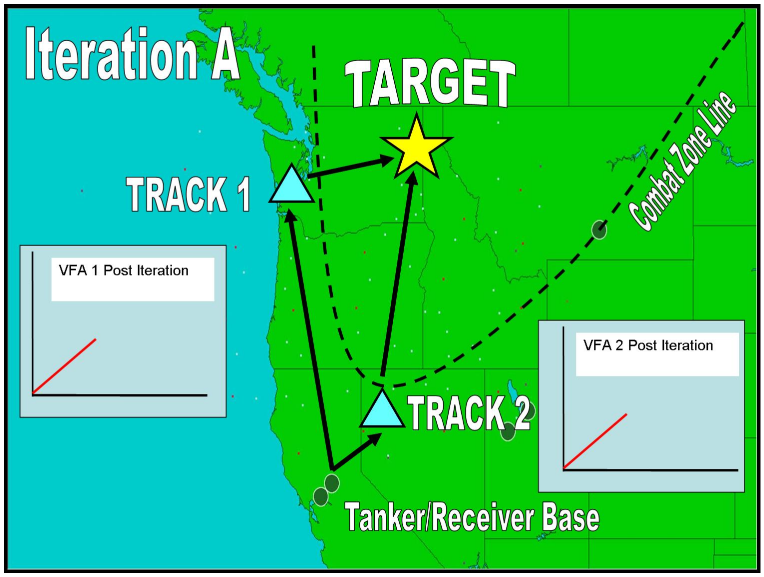

Figure 38: Iteration A - Updating the Value Functions at Both Tracks with No Tankers at either Track

to both tracks since there is a positive value associated with having a single tanker at each track. When the receiver enters the system it is faced with the decision policy that it will travel to the track which has the lowest total distance cost. By setting x arbitrary high the second leg of a receiver mission is much more costly than the first leg when the assignment to track policy is calculated. Therefore, for high enough x the receiver mission will travel to Track 1. With a receiver at Track 1 there is a positive value associated with having a tanker at the track and the value function is updated to show this. Track 2 does not have a receiver and therefore there is no value in having a tanker at the track. The value function at Track 2 is updated through exponential smoothing and the value function reflects the fact that it is less valuable to have a tanker at Track 2.

As the iterations progress and the receiver continually travels to Track 1, the value of sending a tanker to Track 1 continues to remain positive enough to send a tanker to Track 1; however, eventually, the value function at Track 2 will reflect a low enough value that a tanker will not be assigned to to Track 2, as shown in Figure 40.

The previous example illustrates on a small scale how a penalty can induce be-

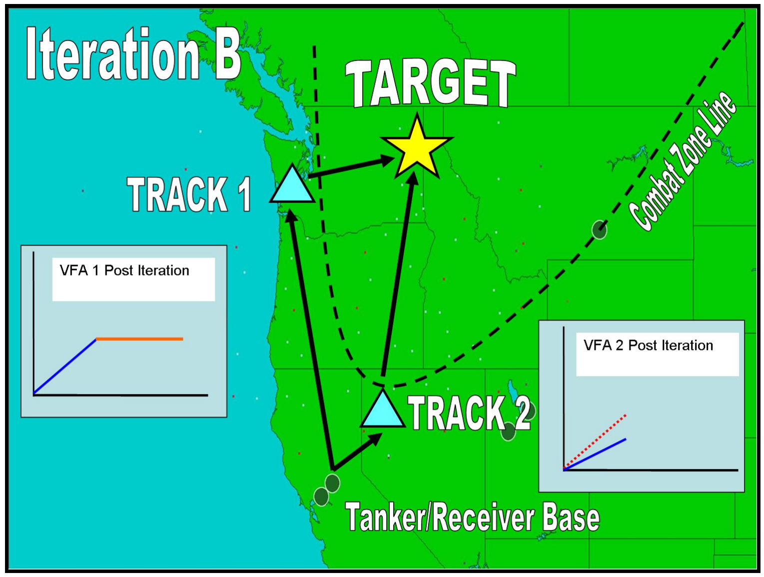

Figure 39: Iteration B - Updating the Value functions at Both Tracks with Tankers at both tracks and receiver at Track 1

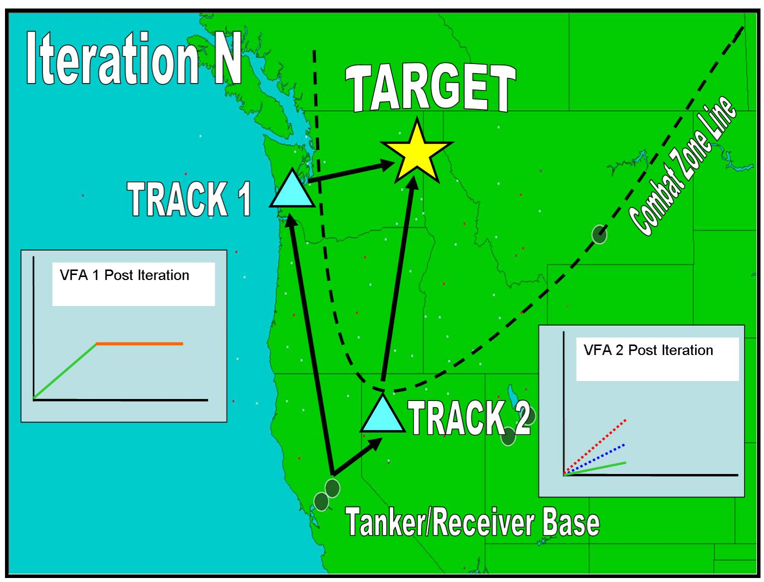

Figure 40: Iteration N - Updating the Value Functions at Both Tracks with a Tanker and receiver at Track 1 no tanker at Track 2

havior which more closely mimics that of real world operational planners. The aerial refueling model optimizes over far more tracks and tankers as well as time periods than the toy example shown above, but the same general framework still applies. The receiver missions are still broken into two distinct parts with the track to target distance holding a greater weight in determining where receivers move than the movement from base to track.

The standard setting used throughout this thesis for receiver "weighting" factor is set at 0.6. When the weighting factor is set to 0.0 the model is indifferent between the relative lengths of the two legs of the trip and merely optimizes both tanker and receiver fuel. As the weighting factor is increased it is expected that the receivers will be refueled closer to their targets. Consequently, as the receiver's movements are more heavily weighted in the model, albeit indirectly, the tanker total fuel cost will stay the same or increase due to the added constraint. The input for the weighting factor, is referred to as the Movement Penalty, shown in table 17

Variable	Iterations	Tankers	Rcvr Penalty	Fuel Ratio	Movement Penalty
Set 1	100	25	10,000	2.18	0.0
Set 2	100	25	10,000	2.18	0.6
Set 3	100	25	10,000	2.18	5.0

Table 17: Large Data Set Inputs Changing Movement Penalty

To measure the changes in the model, the standard approach of looking at the fuel consumption for both the receivers and the tankers is not entirely appropriate. While these measures give meaningful data on the fuel required, there is a more appropriate measure for this series of simulations. For these simulations a measure of the distance the receivers are flying from their tracks to their targets highlights the response of the model to changing the weighting parameter.

The results in Table 18 are illustrated in Figures 41 - 43, which highlight the difference in the distances traveled by the receiver missions in the LDS.

	RcvrFuel	TankerFuel	Delay	MaxDelay	TnkrUsed	Unused	Used
Set 1	1394244	885203	508	11.33	18	0.42	4.25
Set 2	1595082	1525753	486	11.33	20	0.5	4.75
Set 3	2613790	1958372	465	11.33	24	1.58	6.58

Table 18: Large Data Set Outputs After Changing Movement Penalty

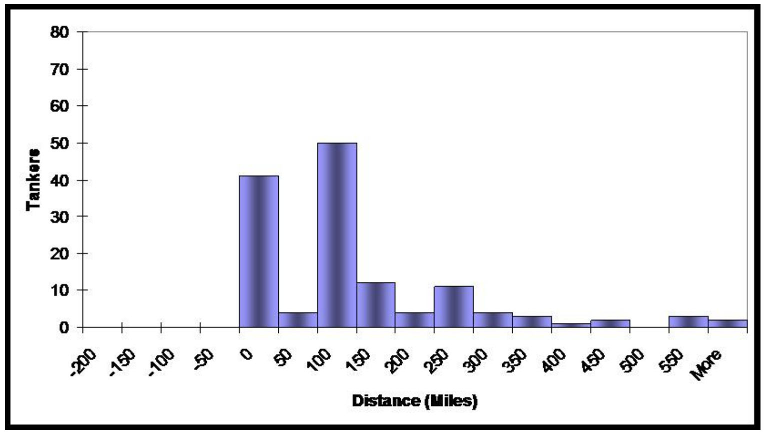

Figure 41: Difference in Track to Target Location for Identical Receivers (Miles)- Movement Penalty Factor 0.0 minus Movement Penalty Factor 5.0

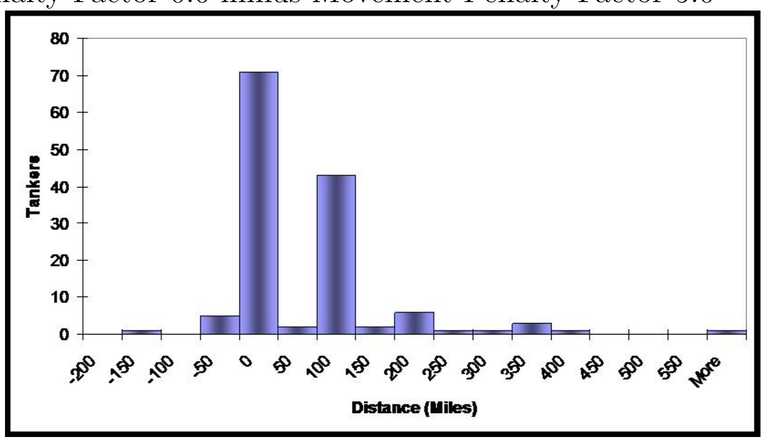

Figure 42: Difference in Track to Target Location for Identical Receivers (Miles)- Movement Penalty Factor 0.6 minus Movement Penalty Factor 5.0

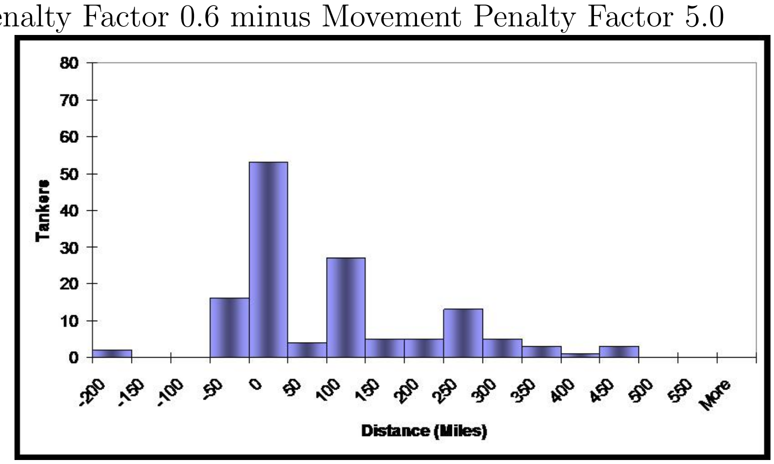

Figure 43: Difference in Track to Target Location for Identical Receivers (Miles)- Movement Penalty Factor 0.6 minus Movement Penalty Factor 0.0

Figures 41 - 43 illustrate the effect on the location of receiver refueling tracks using different penalties. The difference between the distance traveled for identical receivers when there is a penalty factor of 5 versus a penalty factor of 0 is dramatic (Figure 41). The distance is calculated as the lower penalty factor receiver distance minus the higher penalty factor receiver distance so positive values indicate that the lower penalty factor receiver traveled a longer distance. With the higher penalty factor the receivers always fly a shorter distance from track to target for the LDS. The ability to change the behavior of the model so dramatically is an important result for its importance in realistically modeling combat aircraft movements.

During Operation Enduring Freedom in Afghanistan this model could have been particularly useful when examining aerial refueling of US Naval aircraft. During the early stages of OEF, Air Force tankers were based on the island of Diego Garcia and at Romanian air bases, both of which are thousands of miles from the border of Afghanistan. While the tankers were flying in from one location, the United States Navy's aircraft carriers were positioned off the coast of Pakistan in the Indian Ocean. Receivers flying from the aircraft carriers required refueling operations on their way to their targets in Afghanistan. Modeling this problem with the aerial refueling algorithm and the track penalty set to zero, the behavior would likely not be suitable to combat operations as receivers would refuel at tracks which lowered the tankers travel distances. As shown in Figures 41 - 43, when the model is free to optimize without a track to target penalty, the chosen refueling tracks often entail a long track to target distance for the receiver. While the result is mathematically correct, during combat operations the preferred refueling method is that tankers come to a location which is more optimal for the receivers than visa versa. By changing the penalty factors the mission profiles for the OEF missions could be tailored to accurately reflect preferred mission profiles and refuel closer to the targets in Afghanistan then the tanker bases.

Despite the favorable characteristics of the model, a large drawback of assigning a high penalty to the last leg of the receiver missions is that the tanker fuel burn cost incurred increases. Figure 44 illustrates the dramatic increase in the total fuel burned by the tankers when the penalty is increased. The increase in fuel consumption by the tankers as the penalty increases is a direct function of tankers traveling greater distances to tracks which are closer to the receiver's targets. It is interesting to view the Pilotview outputs in Figures 45 - 47, which show how the receiver movements change with the added penalty as well as the differences in the tanker movements in Figures 48 - 50.

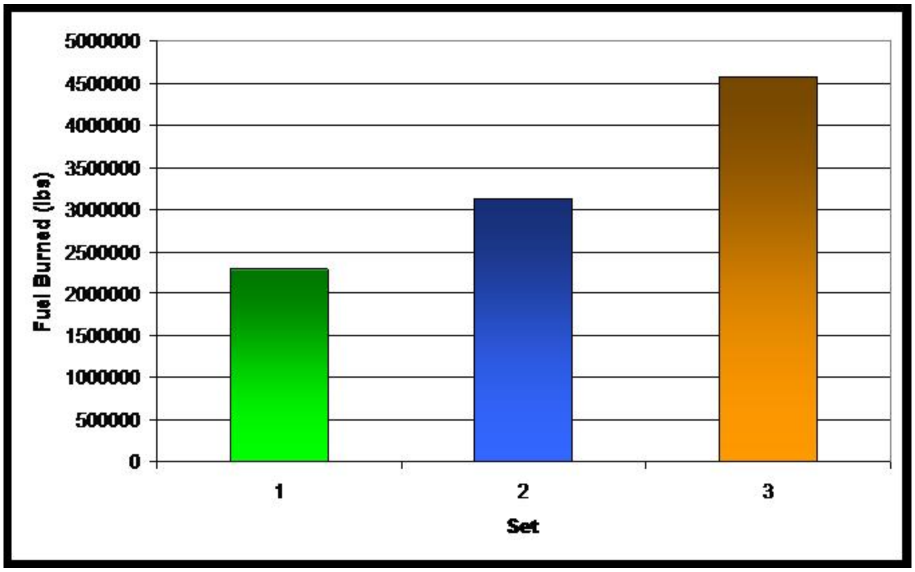

Figure 44: Comparison of Fuel Burned by Set for Varying Movement Penalties - *Set 1* Zero Movement Penalty - *Set 2* 0.6 Movement Penalty - *Set 3* 5.0 Movement Penalty

In the receiver figures, two simulations are overlayed for each time period. The two simulations are with a track to target penalty of 0 and a track to target penalty of 5. Therefore each time period shows the movements of identical receivers through the system. The receiver figures highlight the large differences in the distance traveled between the two simulations. The figures clearly show that when the penalty is set at 5, the distances traveled by the receivers from their refueling tracks to their targets is greatly decreased. An example of this is visible at the top of Figures 46 and 47. At the top of the figures it can be seen that when the penalty is set to 0, the receivers refuel very close to their bases; however, when the penalty is set to 5 the receiver refuels at a track close to its target.

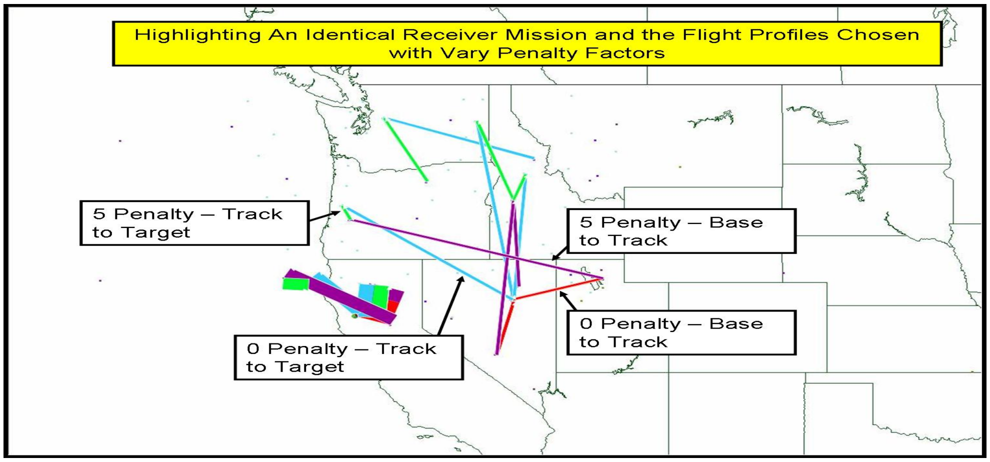

Figure 45: Receiver Movements Comparing 5.0 Movement Penalty and 0.0 Movement Penalty - Time Period 1

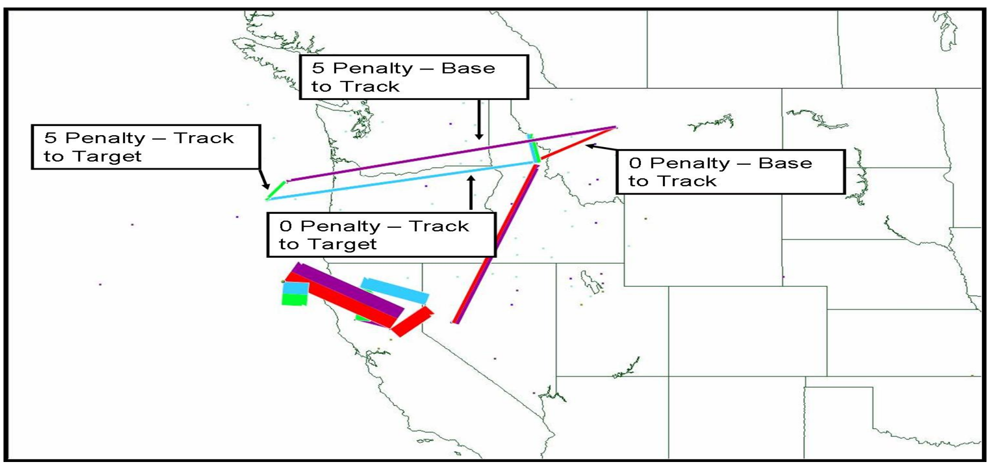

Figure 46: Receiver Movements Comparing 5.0 Movement Penalty and 0.0 Movement Penalty - Time Period 2

Figures 48 - 50 show the movements required by the tankers to refuel the receivers closer to their targets. The figures show two different simulations which are overlayed on the same background. In the tanker example, the tankers are not guaranteed to be identical in each simulation; however, the tankers are refueling identical receiver demands. The interesting aspect of the tanker movements is that in the data set with the high penalty, tankers fly independently across the combat zone. The thickness of the lines represents additional tankers and it can be seen that with zero penalty the tankers tend toward similar tracks. These tracks minimize the tankers total fuel burn

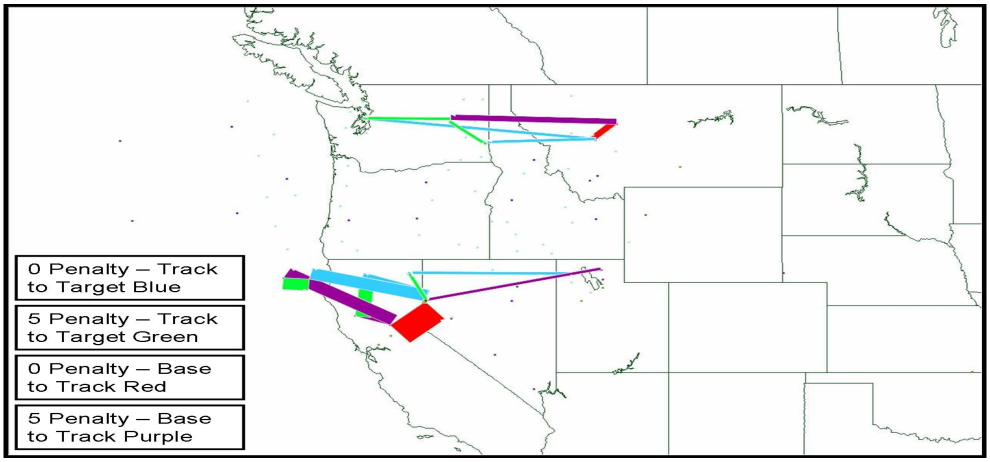

Figure 47: Receiver Movements Comparing 5.0 Movement Penalty and 0.0 Movement Penalty - Time Period 3

since tankers burn fuel at a rate which is more than double that of the receivers. In Figure 50 the differences in the distances traveled by the tankers between the sets is readily apparent and helps to explain the results of Figure 44.

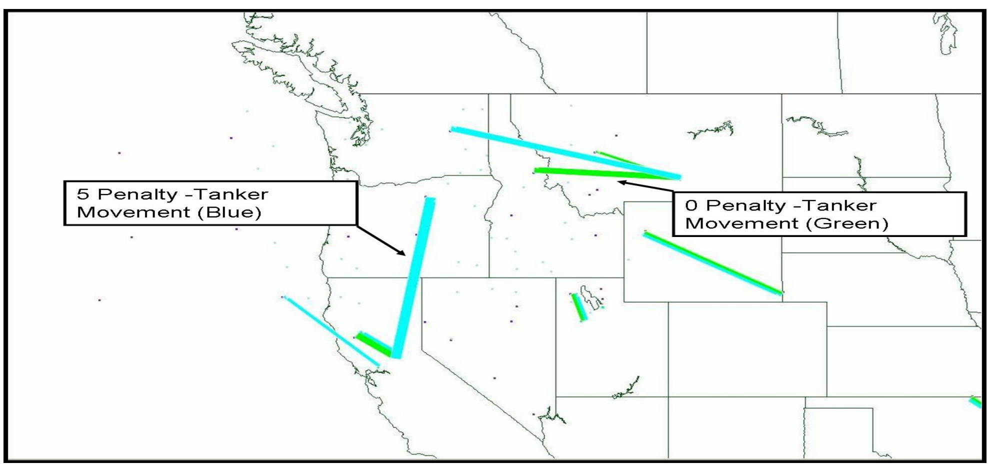

Figure 48: Tanker Movements Comparing 5.0 Movement Penalty and 0.0 Movement Penalty - Time Period 1

The behavior of the model has several advantages and disadvantages which must be weighted in actual combat planning. When the track to target penalty is increased the desired change in the receivers flight patterns is achieved, and they fly to their target with a greater fuel load. The drawback of arriving at their track with a

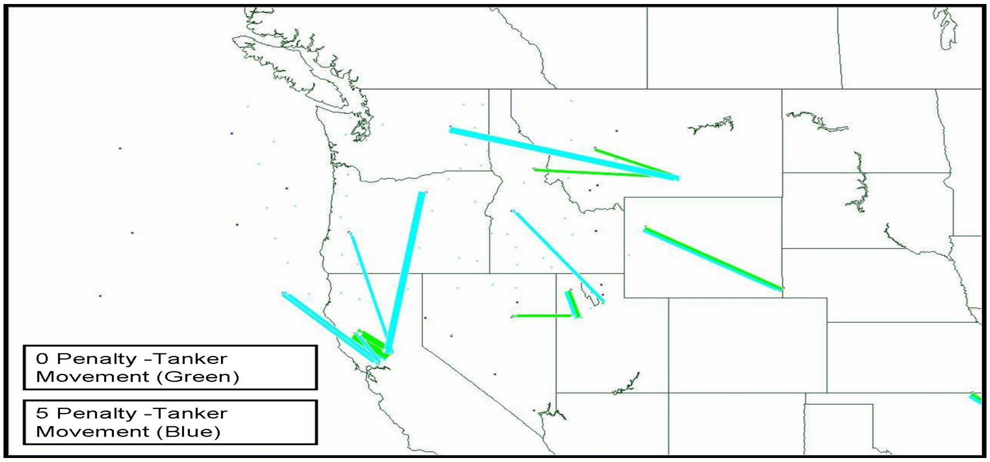

Figure 49: Tanker Movements Comparing 5.0 Movement Penalty and 0.0 Movement Penalty - Time Period 2

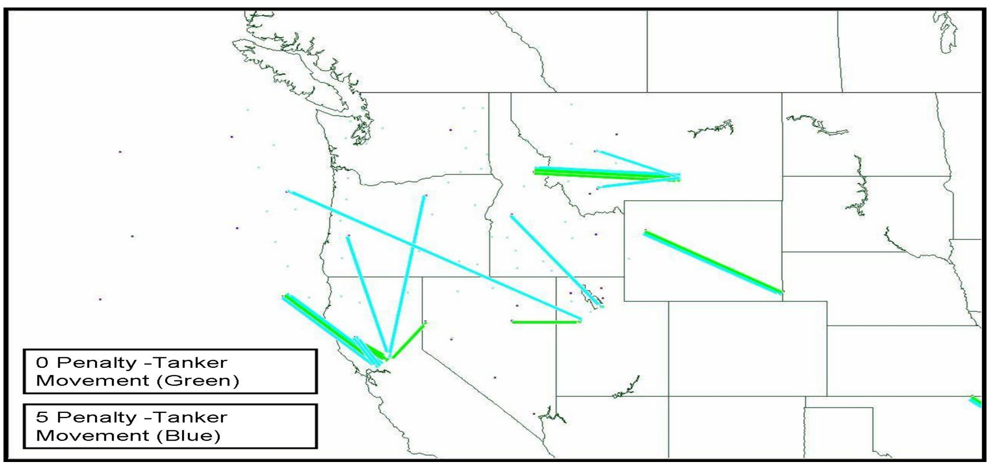

Figure 50: Tanker Movements Comparing 5.0 Movement Penalty and 0.0 Movement Penalty - Time Period 3

greater fuel load is the lack of a common refueling point for receivers. During combat operations tankers have no ability to defend themselves against an enemy attack, and therefore, if they are in a hostile environment they would require fighter escorts to ensure their safety. When the tankers are all located at common refueling tracks it is easier to protect the airspace around the refueling zone than if there are many tankers spread around the combat zone. Therefore, in a time of insecurity early in a conflict when air superiority is still contended, it might be preferable to have common tanker refueling points. A major strong point to this model is its ability to produce both types of receiver/tanker mission profiles with detailed outputs which can guide the combat planner's decision making process.

4 *Extensions - Changing Inputs and Stochastic Demands*

The following sections examine several aspects of the model which do not involve changing parameters within the model. Rather, a series of tests on the adaptability and robustness of the model are shown. The tests focus on introducing stochastic demands of varying types, which include varying receiver arrival times, receiver mission fuel demands, receiver mission loads within the system, and the ability of the model to solve perturbed inputs. In addition to showing the robust nature of approximate dynamic programming, the following sections provide insight into how a mission planner could exploit the model's attributes for specific types of data sets. The following tests show the general nature of solutions as well as the adaptability of the model to changing inputs, which is important when planning for uncertainty such as in aerial refueling.

4.1 *Using Results to Guide Inputs - Stochastically Perturbing Refueling Times*

The solutions illustrated throughout this thesis have all been generated from on a static data set. During a simulation the algorithm has seen identical receiver demands in each iteration and created value functions which guided tanker and receiver movements. These solutions have been appropriate for combat planning purposes, and we would expect that they would work in real world applications as they are identical to the current solutions that also use static data sets. However, in the application of the solutions to the real world, one could expect that receivers are not identical to the projected receivers and that the receivers arrive 10 minutes early or late or that their fuel levels vary from the projected fuel levels initially planned. For a model to be successful in real world applications it must be able to absorb the stochastic nature of the real world without the solution imploding, which in the aerial refueling problem would be realized through planes falling out of the sky (not a good way to test a solution).

Within the aerial refueling simulator, the ability to adapt to uncertainty has been hidden in plain sight. The statistic which shows how well the model can adapt to varying refueling times and refueling loads is the fueling delay statistic. The fueling delay shows how long planes are expected to wait in a queue for a tanker after their planned refueling time. In the model the fueling delay given for each plane illustrates how well that plane could react to changes within the system. A plane with no fueling delay is not required to wait for refueling since it is assigned to a tanker with no queue, or it is the first plane in the queue. A plane with a long fueling delay is required to wait in a queue for an extended period of time as it is either in a queue with a large number of receivers or in a queue behind a receiver which requires a large fuel offload. For a model to stand up to the actualities of aerial refueling it is required that there exist very low fueling delays for each receiver. Since mission planners usually do not tax the safety reserves of planes requiring refueling, it is clear that for a receiver with a low fueling delay a sufficient fuel reserve must exist to absorb any uncertainties of the system. Figure 51 shows that the fueling delays are modest for the base LDS simulation, with a maximum value of 14.33 minutes. In this model the expectation is that variations in the refueling times and arrival times would not cause the planes to fall out of the sky as each plane is not delayed for an extended period. Additionally, after the aerial refueling problem has been solved, the mission planners could easily adjust the expected arrival times of receivers within a few minutes to decrease any long queuing within the system.

When receivers are delayed for a short time interval, it is usually because two or more identical receivers arrive at a track location at the same time. When multiple receivers arrive at a track at the same time it is often less costly to refuel both of them with one tanker, causing a queue, than to move in another tanker to eliminate queuing. The data sets are constructed in such a way that there are many instances of multiple receivers being clones of another receiver mission and therefore arriving to a track at the same time. The cloned receiver missions are illustrative of fighters flying in pairs to a target or a fighter escorting a bomber to a target, which occurs in

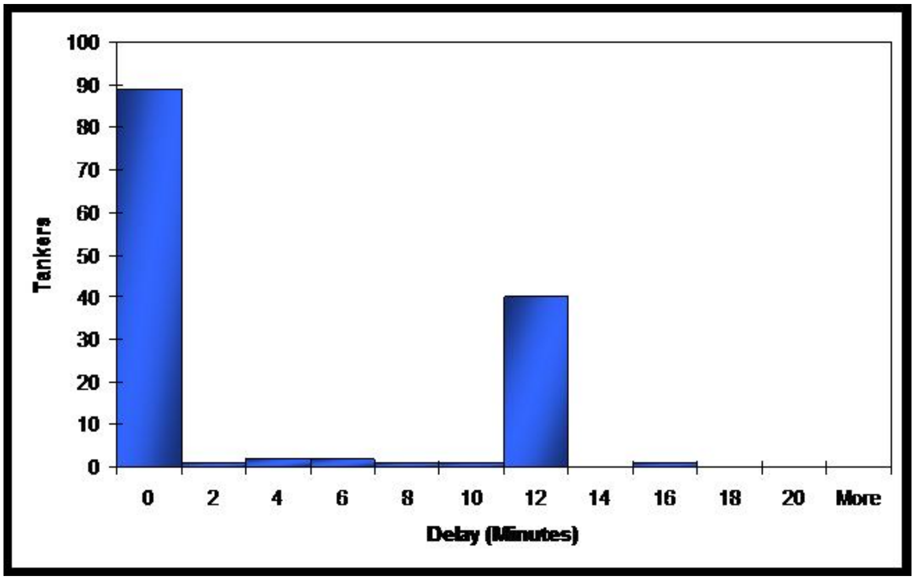

Figure 51: LDS Fueling Delay Base Case

actual mission planning. The important aspect of modeling pairs of receivers flying a common flight plan is that they both arrive at the target area at the same time. While the data sets are constructed to have receivers refuel at identical times in practice it is not necessary that the receivers refuel at identical times. It is important that the receivers refuel at the same location and similar times; however, the overwhelming concern is that they arrive on target together. Additionally, it is often not reasonable to assume that receivers have identical launch times and therefore refueling times if they are both taking off from an aircraft carrier. Thus slightly perturbing refueling times is not an unreasonable compromise of the data set for the goal of reducing queuing within the system.

A mission planner who has run a data set and found queuing times to be unacceptable for identical pairs of receivers could alter the refueling times to lessen the queuing. After examining the initial results from the LDS base simulation, a mission planner could stagger refueling times slightly for identical receivers. A change in the refueling times for identical receivers would be expected to reduce queuing time and allow for greater variability in the process of refueling, without changing the goals and capabilities of the mission profiles.

This is a reasonable goal of a mission planner and is easily implemented through changing refueling times slightly and rerunning the model. To implement the changes in receiver refueling times, a mission planner could go through the missions and manually change the refueling times; however, in a large data set accomplishing this goal could be a long procedure. Instead of manually shifting refueling times, the model was set to introduce randomness into the refueling times. For the base LDS all inputs are deterministic so every simulation produces identical results. To change the refueling times, when the deterministic refueling times were read into the system they were perturbed. The perturbation used a random number generator from a fixed interval to add between [-10, 10] minutes to each receiver mission. By shifting the receiver missions, the model was able to eliminate identical refueling times.

A series of five simulations with perturbed refueling times were run. All five simulations showed a decrease in queuing times, which was a direct result of receivers not having identical refueling times. To account for the stochastic nature of the new data sets when reporting the results the five perturbed solutions are averaged. In the base LDS a pair of identical receivers which are refueled by the same tanker would accrue large queuing cost. In the perturbed LDS the same "identical" missions now come to the refueling track at slightly different times, and therefore while they are still refueled by the same tanker they are not forced to wait in a queue for as long as the base case.

As shown in Figure 52, when the mission planner varies the refueling times of the receivers slightly, the results are very similar to the base case with respect to the total cost of the system; however, as shown in Figure 53 the fueling delays are decreased dramatically. The reduction in fueling delays gives the model the flexibility to absorb the uncertainties of the real world to a greater degree, and is accomplished without changing the ability of the receivers to complete their initial missions.

The success of introducing slight perturbations into refueling times and dramatically reducing queuing in the system is a strength of the model. The small shifts in refueling times do not dramatically influence the decisions within the system; how-

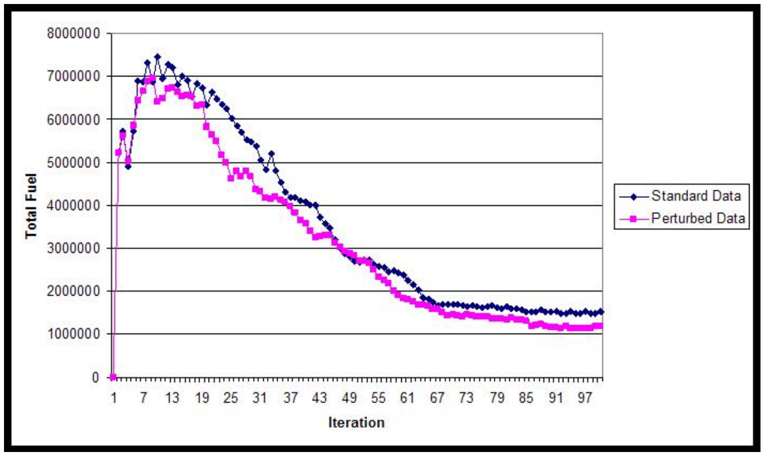

Figure 52: Total Cost for the Base LDS Simulation and the Compiled Perturbed Refueling Time Simulations

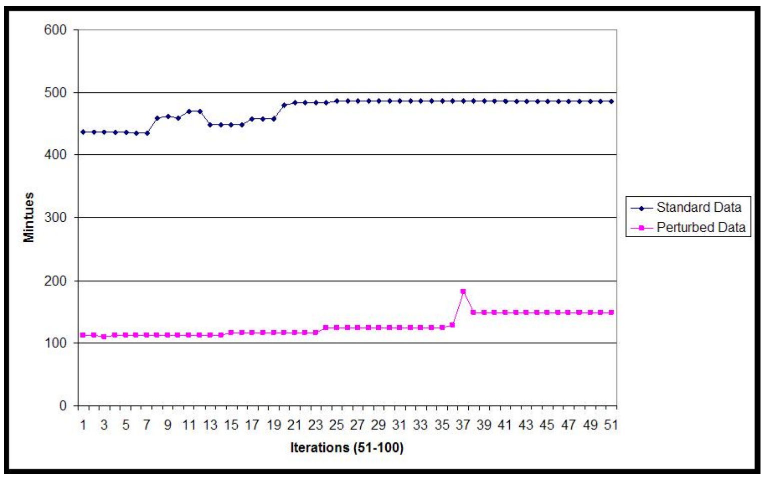

Figure 53: Fueling Delay for the Base LDS Simulation and the Compiled Perturbed Refueling Time Simulations - Iterations 61 - 100

ever, they greatly reduce queuing. The results shown by perturbing the refueling times also illustrate the flexibility of the initial solution for the base LDS simulation. The base LDS simulation had many "identical" receivers; however, in practice one receiver would arrive slightly before or after their counterpart which would lead to decreased queuing. The perturbations to the refueling times shown illustrate how well the base simulation would be able to handle the stochastic nature of aerial refueling. This result shows that the aerial refueling model is very robust for varying refueling times and the base results are stable enough to handle the actual aerial refueling operations.

4.2 *Stochastically Varying Fuel Demands*

The current model employs a predetermined fuel offload for each receiver mission. While it is reasonable for modeling to assume that receiver missions require a fixed fuel level, a positive attribute of the model would be an ability to accommodate varying fuel levels. When increasing the stress on the model through stochastic fuel demands it is hoped that a variety of poor results are not induced, such as: increased fueling delays, mission failures, or tankers running out of fuel.

To test the ability of the model to respond to stochastic fuel levels, two different types of simulations were run. The base simulation (deterministic) took the SDS and looped over the missions, increasing the fuel demands by 20 percent over the original fuel demand for 50 percent of the missions.

$$\left(\; FuelDemandReceivers \; \right) \;\; = \;\; \sum_{j \in \mathcal{J}} 1.2_{(\tilde{P}(>.5))}(FuelDemand_j) + 1.0_{(\tilde{P}(<.5))}(FuelDemand_j)$$

The new data set, $SDS2050$, was optimized for twenty iterations up until a stopping iteration n_u. After twenty iterations the value function approximations (VFA) were fixed and a new input data set was tested on the trained VFA. The new data set, $SDS2050_i$, was identical to $SDS2050$ except that the fueling demands were perturbed. For each deterministic data set and its associated VFAs, ten perturbed data sets were tested. In this manner the ability of deterministically trained VFAs to optimize perturbed data sets were tested. Since each set of deterministically trained VFAs is only one sample realization (the sample path Ω is simply a series of identical ω_i), 15 different simulations with different original $SDS2050$ were run to find the average ability of the data sets to optimize the stochastic data sets, $SDS2050_i$.

The counterparts to the deterministically trained VFAs are stochastically trained VFAs, which are created through changing the input data set at each iteration of the VFA training phase. While the deterministically trained data simulations take a sample realization and optimize on the single realization for 20 iterations, the stochastically trained simulations use a different sample realization for each iteration. Therefore, the model is constantly adjusting to optimize VFAs with changing demands and the sample path, Ω, is responsive to both ω_i and the ordering of the realizations. As with the deterministically trained VFAs, the stochastically trained VFAs are trained until n_u and then the trained VFAs were tested with ten stochastic data sets $SDS2050_i$. The updated algorithm for incorporating both stochastic data sets as well as stopping the updating of value functions is shown in Figure 54.

Step 0: Initialization:

Step 0a. Initialize $\bar{V}_t^0, \quad t \in \mathcal{T}$.

Step 0b. Set $n = 1$.

Step 0c. Initialize R_0^x (The set of all tankers in the system).

Step 1: Choose a sample realization ω^n if deterministic run and n = 1, or if deterministic run and $n > n_u$, or if stochastic run. For $t = 1, 2, \ldots, T$. (Standard receiver missions with altered fuel demands) do:

Step 2a: Create the linear program from the available tankers and associated value function approximations:

Step 2b: Solve the optimization problem:

$$\max_{x_t \in X_t^n} [(C_t(R_t^n, x_t) + \bar{V}_t^{n-1}(R^{M,x}(R_t^n, x_t))]$$

Step 2c: Simulate the receiver refueling and queuing to find $\hat{v}_t^n(R_t^x)$

Step 2b: Increment $R_t^x \pm \epsilon$, at all tracks.

Step 2d: Re simulate the queues with the $\pm\,\epsilon$ to find the derivatives which are $\hat{v}_t^n(R_t^x(\pm\epsilon))$

Step 2e: If $t > 0$ and $n < n_u$ (Where n_u is a predetermined iteration for stopping updates) Update the appropriate value function using:

$$\bar{v}^n(r) \quad = \quad \begin{cases} (1 - \alpha_{n-1})\bar{v}_{t-1,a}^{n-1} + \alpha_{n-1}\hat{v}_{ta}^n & \text{if } r = R^n \\ \bar{v}^{n-1}(r) & \text{otherwise} \end{cases}$$

Step 2f: Update the States:

$S_{t+1}^n = S^{M,W}(S_t^{x,n}, D_{t+1}, W_t)$

Step 3. Increment n. If $n \leq N$ go to **step 1**.

Step 4: Return the value functions, $\{\bar{V}_t^n, t = 1, \ldots, T, a \in \mathcal{A}\}$.

Figure 54: An approximate dynamic programming algorithm to solve the aerial refueling problem incorporating stochastic data sets.

To create meaningful results when testing stochastic data, the data sets are averaged so that conclusions are not drawn from a single sample path. For both the stochastic and deterministic data sets 15 separate simulations were run and the results were compiled.

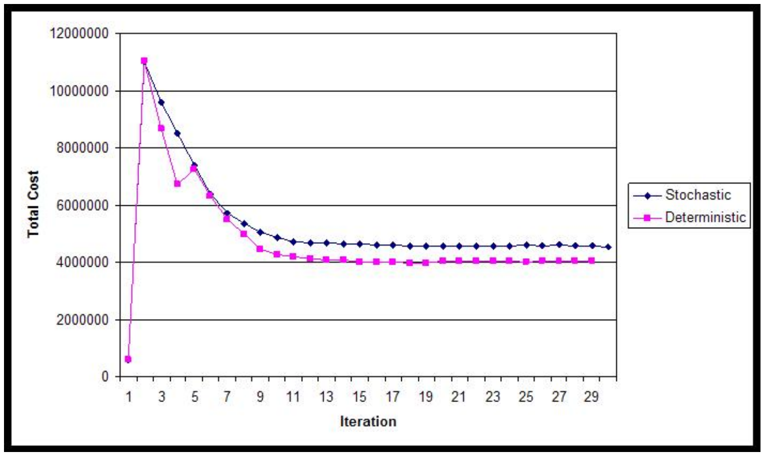

Figure 55: Total Cost Stochastically Trained Simulations versus Deterministically Trained Simulations - Training for 20 iterations and Testing over the last 10 iterations

Since there is a high cost associated with long fueling delays and mission failures, the expectation is that stochastically trained simulations will send out more tankers during its training phase than the deterministically trained simulations. As shown in Figure 55, during the twenty training iterations the stochastically trained total cost is higher than deterministically trained simulations. The components of the higher cost are the total fuel burn by the receivers as well as the tankers. The higher fuel burn of the receivers is caused by a greater amount of queuing in the system (Figure 56), as the system cannot optimize the tanker fleet as precisely as in the deterministic. The second component of the increased cost is contributed by the increased tanker fuel cost (Figure 57). The increase in the tanker cost is due to the system sending out additional tankers in the stochastic simulations due to the increased value of tankers at tracks when the demand is not as clearly known.

The results during the training phases between the two simulations are intuitive and mirror the decision a person would likely choose. When a mission planner is given

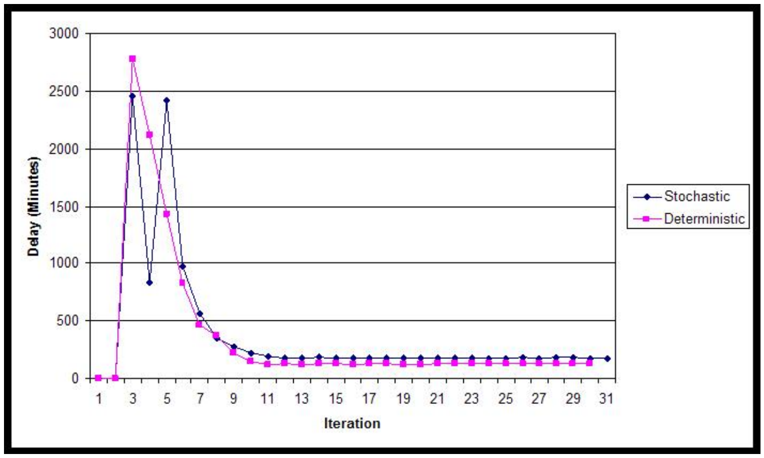

Figure 56: Fueling Delay Stochastically Trained Simulations versus Deterministically Trained Simulations - Training for 20 iterations and Testing Over the Last 10 Iterations

uncertainty he would likely err on the side of caution and place additional tankers in the sky to limit negative outcomes. This is the behavior shown during the training iterations when the model has an approximation of the future demands and sends out additional tankers to limit excessive fueling delays and mission failures.

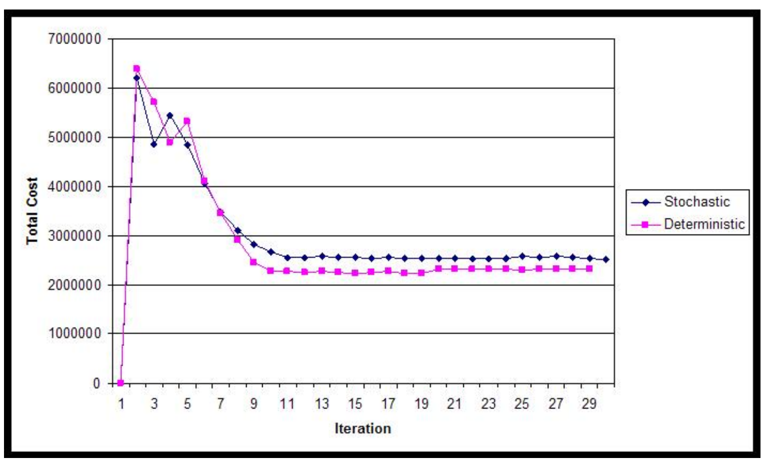

Figure 57: Tanker Cost Stochastically Trained Simulations versus Deterministically Trained Simulations - Training for 20 iterations and Testing Over the Last 10 Iterations

The testing phase on the trained VFAs is also instructive in that the output data does not shift any appreciable degree. The a priori expectation is that the stochastic simulations would create VFAs which optimize better during the testing phase than the deterministically trained VFAs. This expectation is based on the fact that the stochastic VFAs are more general and value having more tankers at tracks to accommodate perturbations in fuel demands than deterministically trained VFAs.

However, the results showed that the deterministically trained VFAs are general enough to accommodate the instability in fuel demands. The stochastically trained VFAs also perform well when tested, but the excess tanker movements dictated by the VFAs do not improve the total receiver fuel burn or fueling delay. The results, while unexpected, illustrate that the VFAs as constructed can handle significant perturbations to the receiver missions fuel levels. While the perturbations to the fuel levels are significant, they represent a small cost within the system. Increasing a fuel demand from 20,000 lb to 24,000 lb (which is an average receiver mission) only increases the fueling time by a few minutes, and therefore any planes queuing behind that plane will only encounter a few extra minutes of queuing. This small increase in queuing results in the system accruing a very small change in total cost. Where the fuel load is increased a great deal, such as an offload to an EP-3 from 100,000lb to 120,000lb, it occurs with tankers which have no associated queue since the original offload exhausts most of the tankers fuel. The added cost of the system thus does not significantly change the results of the model.

The results shown by the deterministic data set's ability to handle stochastic fuel levels once again illustrates the robust nature of the aerial refueling model. The ability of the model to assimilate varying receiver refueling times as shown in Section 4.1, as well as varying fuel levels, shows that a deterministically trained data set's solutions are very flexible. Mission planners want aerial refueling solutions which are both efficient and reliable in the real world and the aerial refueling model meets both of those objectives. In the following section, the VFAs will be tested with much greater perturbations to the system as the number of receivers will vary throughout the simulation. This test will go beyond the expectations of mission planners and again illustrate the robust nature of the aerial refueling model.

4.3 *Receivers Everywhere!!(Modeling Varying Receiver Demands)*

The method of Approximate Dynamic Programming is a very powerful approach when applied to stochastic demands since it can build value functions which account for the varying demand levels. A standard example of the use of ADP with stochastic demands is illustrated throughout Powell's text (17) in the nomadic trucker example.

In the nomadic trucker example, at each time period and location a load with a certain value to be carried to a new location can exist or not exist. If the trucker is at that location then he observes the value of being at that location at that point in time. If the trucker is not at that location then he never observes the load and it is assumed to disappear (another trucker moves the load). Within the nomadic trucker example, it is easy to implement stochastic demands since if a load is not carried there is not a downside other than lost revenue since the demand leaves the system. Therefore, over a simulation run a trucker can periodically sample locations and find an approximation of the value of being at locations at a certain times. To scale up the nomadic trucker example, if you assume that it is a trucking company and they can send multiple trucks to many locations (as is the case with the aerial refueling model) then the model resembles the aerial refueling model. In the larger trucker model during the simulation the company might find that on Tuesday mornings it is optimal to have four trucks in Miami since they expect four loads. If on Tuesday morning three loads appear, then the company has no problem and has merely wasted a resource that might have been able to fill a demand elsewhere. If instead on that Tuesday there are five loads then the company moves the four loads and ignores the fifth load. In both of these examples the trucking company would update their estimation of the value of a having four trucks in Miami on Tuesday morning, but the company would not drastically alter the number of trucks they send to Miami.

The aerial refueling constraints are much different since within the system unsatisfied demands do not disappear from the system. The aerial refueling model is similar to the trucking company with multiple trucks in that if it has too many tankers at a location with few demands it will decrease its estimation of the tankers required. The large difference between the two models is when the aerial refueling model has too few tankers to fulfill the receiver demands. The receiver demands do not disappear from the system, rather, large penalties for refueling delays and receiver crashes accrue in the system. It is the large penalties associated with receivers crashing which help to drive receiver mission failures to zero in the initial iterations of the model, but they can also limit how effective the model is at handling stochastic demands.

While the nomadic trucker example does not require any structure to the demands entering the system outside of a distribution of demands, this is not the case for the aerial refueling model. The aerial refueling model cannot handle a series of random missions at each iteration due to the large penalties which accrue in the system. Therefore, the randomness of the missions must be limited to provide a measure of stability to the system. With the need for stability in mind, an existing data set, SDS, provided the foundation for the stochastic data set. From the SDS the receiver missions (demands) in the system are randomly sampled for each iteration. Given the structure and sampling of the new data set, the dynamics of the system are not radically altered but the ability of the model to incorporate new information at each iteration is illustrated.

4.3.1 *Simulation Set Up*

The structure of the stochastic and deterministic simulations are similar to that of the stochastic fuel levels section (4.2); however, a brief summary is provided for this specific simulation.

To test the ability of the model to incorporate a random sampling of receiver missions, the simulations were broken into two phases. The first phase of the simulation was the "training" phase in which the model operated in its normal mode and updated the value functions after every iteration. To train the value functions and then test their ability to incorporate stochastic data, the value functions were trained on both a deterministic data set and a stochastic data set. For the deterministic data set in the first iteration, a random subset of the receiver missions was chosen and used to train the value functions. In choosing the receiver missions which would enter the system, the formula below was used which looped over all available receiver missions, $\mathcal{J}$, and entered them into the system using an indicator function.

$$\left(\ Receivers \ \right) = \sum_{j \in \mathcal{J}} j * 1_{(\tilde{P}(>.8))} \tag{31}$$

Therefore, in each deterministic simulation the receiver missions entered in the model were different sample realizations; however, the sample paths for each simulation were fixed throughout the training phase. To train the model with the stochastic data, the receiver missions which entered in the model were changed before each iteration, again by Equation 31. In this sense the sample path seen by the stochastic training simulation was much more complex than that seen by the deterministic training simulation. The sample path for the deterministic training model was determined at the beginning of the simulation and was only concerned with the number of receivers entered into the system. For the stochastic training model the sample path concerned a different sample realization at each iteration, and therefore both the number of receivers entered into the system as well as the timing of the receivers entering into the system added randomness to the model. This is a fairly extreme way to test the value functions, but it helps to show the stability of the system and its applicability to real world situations.

After the training phase for both the stochastic and the deterministic simulations, the value functions were frozen at their current values and then the stability of the value functions was tested. To test the stability of the value functions at each iteration of the testing phase, a different sample realization of the receiver missions was run through the model using the fixed value function approximations to guide the movements of the tankers in the system. The sample realizations were again a subset of the SDS which was constructed using Equation 31. Since the receiver missions are pulled from an existing data set, the expectation is that the stochastically trained simulations will be able to incorporate the stochastic sample realizations of the testing phase better than the deterministically trained runs.

Since each run of the model for both the stochastic and deterministic training runs followed different sample paths, the results for 15 simulation runs were aggregated to find how well on average both systems worked. Fifteen runs were used due to the apparent stability of the averages after 10 simulations and a the desire to build in a buffer. While it is entirely likely that given a different set of 15 runs the results would be different, the results from this test were stable, and therefore conclusions drawn about the model would not differ to any appreciable degree.

4.3.2 *Results*

Since each simulation was split in two distinct phases, training and testing, the results of each part are examined separately. The training phase for both the deterministic and stochastic data sets was run for 19 iterations, and the testing phase was the following ten iterations. During the training phase, shown in Figure 58, the model optimizes behavior for both the deterministic data sets as well as the stochastic data sets.

The major difference between the simulations is that the deterministic optimization is much smoother and lower than that of the stochastic optimization. This result is expected since in the deterministic simulations the model saw identical sample realizations for all 19 training iterations, while in the stochastic simulations each iteration saw a different sample realization. While the total fuel used in the stochastic simulations was higher than that of the deterministic simulations, an interesting result about the fueling delays in the training phase emerged which is shown in Figure 60. The increased delay for the deterministic simulation accounts for a huge increase in total receiver fuel burn which is shown in Figure 59 and discussed further throughout this section.

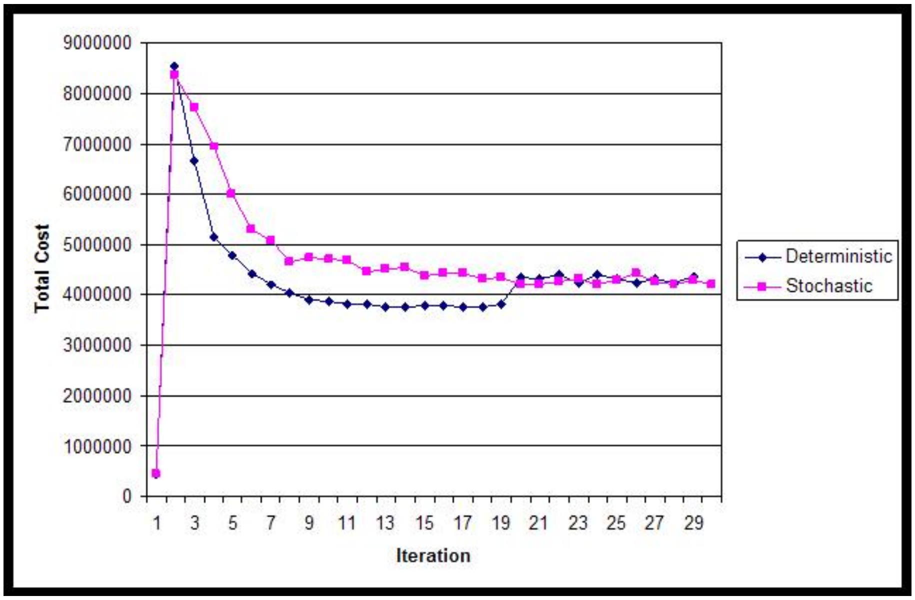

Figure 58: Total Cost Stochastically Trained Simulations versus Deterministically Trained Simulations - Training for 19 iterations and Testing Over the Last 10 Iterations

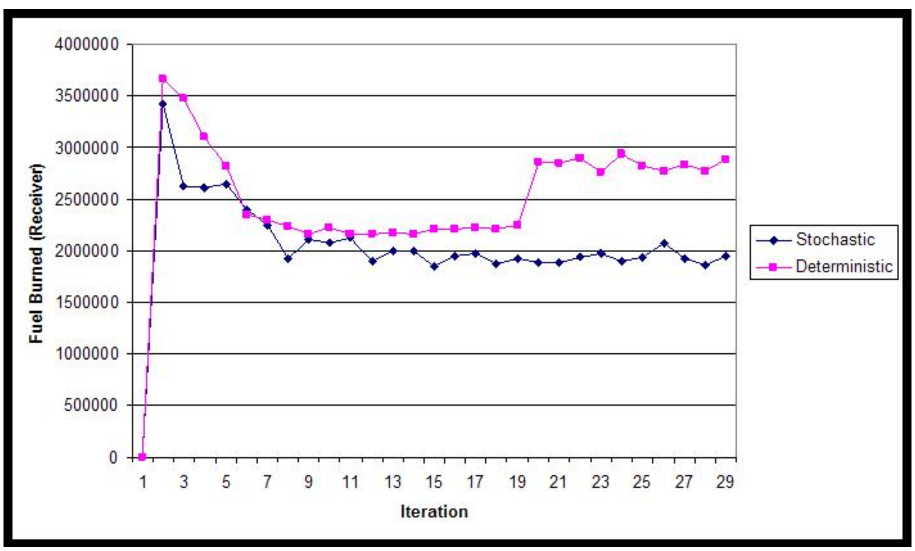

Figure 59: Total Receiver Fuel Burned Stochastically Trained Simulations versus Deterministically Trained Simulations - Training for 19 iterations and Testing Over the Last 10 Iterations

Since the deterministic data sets see the same receiver missions in each iteration it is expected that the deterministic data simulations would have a lower fueling delay than the stochastic simulations. The result which is opposite of the expectation, is

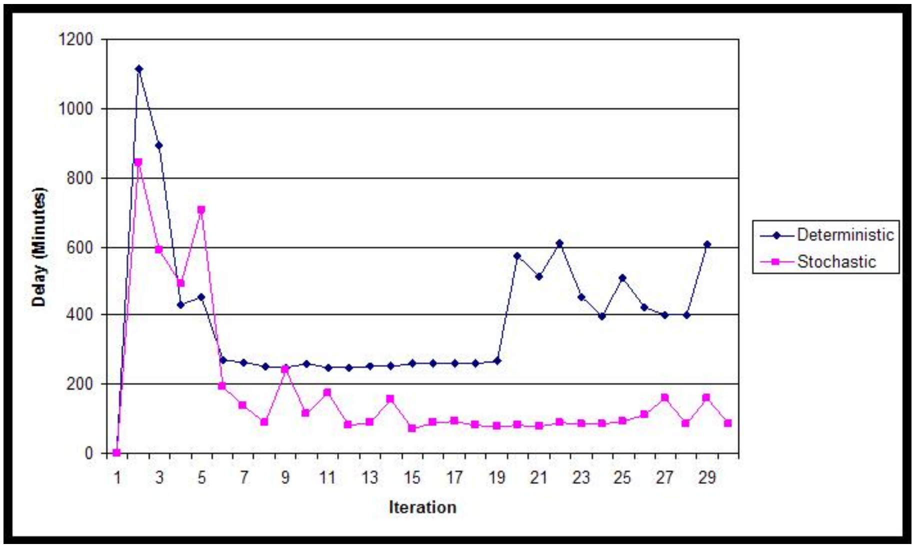

Figure 60: Total Delay Stochastically Trained Simulations versus Deterministically Trained Simulations : Set 2 - Stochastically trained fuel demand : Set 1 - Deterministically trained fuel demand

not a shortcoming of the model, but rather an illustration of how the model views queuing time and tanker movements. Within the model, as mentioned earlier in this thesis, there is a changeable parameter which concerns the amount of delay a receiver can accommodate before a major negative penalty is accrued. For the aerial refueling model simulations this parameter was set at 15 minutes which allowed for queuing to occur in the system. If the parameter was set to zero minutes, then the model would see no reason to have planes wait in a queue, and instead of having a tanker refuel several receivers back to back, each receiver would be refueled by its own tanker. Obviously, the former behavior of queuing is preferable to the latter, and hence the parameter is set at 15 minutes. In a deterministic simulation the model attempts to minimize the queuing time of each receiver, subject to the goal that fueling delay is less than fifteen minutes. When the queuing time is under fifteen minutes the fuel burn rate of a receiver is far less costly than sending out an additional tanker, and thus in a deterministic model there are many receivers which queue between zero and fourteen minutes.

The stochastic data simulations are also bound by the same parameter; however, unlike the deterministic simulations the stochastic simulations do not know which missions will be in the next iteration. Given that limitation how do the stochastic simulations keep fueling delays under 15 minutes, by sending out as many tankers to a locations as possible. Since all of the samples are drawn from the SDS over a series of iterations, each available receiver mission is likely to be seen within the system. If a tanker is unavailable for a receiver at that time and the mission fails, or there is a large fueling delay, then the value function approximations respond by putting a high value of having additional tankers at that track within that time period. The model learns quickly to send an overabundance of tankers to locations to mitigate possible mission failures and fueling delays. As shown in Figure 61, the stochastic simulations use far greater tankers per time step than the deterministic simulations. On average throughout the simulations of the available 40 tankers in the system, the deterministic simulations set used 16 tankers while the stochastic simulations used 25 tankers.

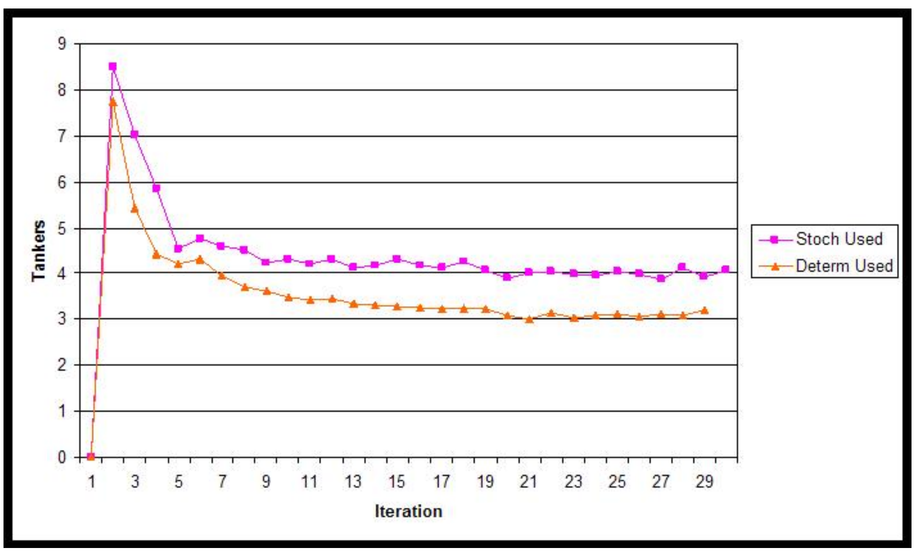

Figure 61: Tanker Usage Per Time Step Stochastically Trained Simulations versus Deterministically Trained Simulations - Training for 19 iterations and Testing Over the Last 10 Iterations

It is interesting to note the differences between the training phases of the simulations; however, these simulations were run to test the differences in the stability of the trained value functions when facing stochastic data sets. The expectation is that while the deterministic simulations excelled in reducing total cost, the value functions will not be able to accommodate stochastic data as well as the stochastically trained value functions. For both the stochastic and deterministic simulations, the trained value functions were tested with 10 different sample realizations of receiver missions. Neither the stochastic nor deterministic simulations' value functions were updated during the testing, but rather it was a test of how flexible the value functions were in accommodating different demands. Looking again at Figure 55, each of the last 10 data points are averages across all fifteen simulations at that iteration. Therefore, while it is useful to see the total cost plotted as iterations, there is no reason to compare Iteration 23 from the deterministic simulation with Iteration 23 from the stochastic simulation. For Figure 55, you can see that it appears as though both the stochastic and the deterministic simulations optimize equally during the stochastic testing. As shown in Figure 62, which is the average across all 150 sample realizations from both the deterministic and stochastic simulations, the difference between the two is only 55,468 pounds of fuel (.01 percent). The differences between the simulations appear to be smaller than the breadth of a single hair. However, while the total cost are similar it is instructive to examine the components of the total cost.

The two components of the total cost are the total receiver fuel cost and the total tanker fuel cost. Looking again at Figure 61, it is obvious that the stochastic simulation will have a much greater tanker fuel cost due to it sending more tankers. However looking at Figure 59, it is obvious that the receiver fuel cost is much lower for the stochastic simulation than the deterministic simulation due to much less queuing. The reason for this is the ability of the stochastically trained simulations to accommodate stochastic receiver missions and maintain a low overall fueling delay in the testing phase. The deterministically trained value functions cannot readily handle the stochastic receiver demands and the fueling delays go through the roof. The fueling delays for the deterministically trained simulations are almost five times those of the stochastically trained data sets.

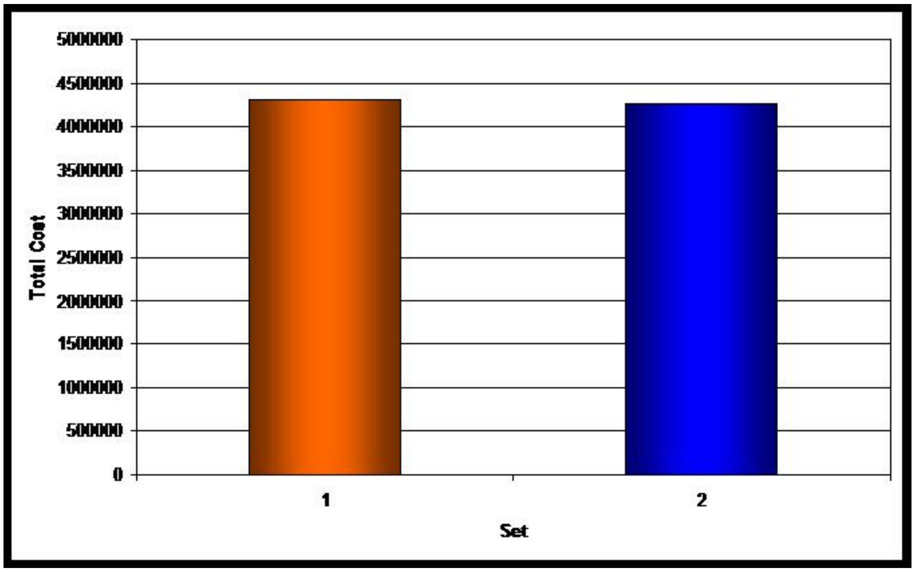

Figure 62: Total Delay Deterministically Trained (Set 1) versus Stochastically Trained (Set 2): The Testing Phase

The conclusions from these simulations are not as readily apparent as anticipated; however, they do illustrate both the technical and the subjective stability of the value functions. The stability of the value functions and their ability to respond to stochastic data are shown through the lack of variability when the stochastically trained value functions were tested on a stochastic data sets, especially when compared to the huge cost increase of the deterministically trained value functions. While it would have been a bonus to see a great total cost difference during the testing phase, the more important result was the differences in the stability of the solution and this showed that the value functions of a stochastically trained simulation are more stable than a deterministically trained simulation as expected. The subjective conclusions from these simulations focus on the preferences of mission planners to minimize fueling delays, particularly fueling delays longer than a preset time. The stochastic simulations were far and away the better choice when measuring fueling delays and may be useful for Air Force mission planners. While changing the entire composition of the receiver missions between simulations is not likely to benefit mission planners a great deal, the model can incorporate such uncertainties. More likely mission planners who know a base set of missions, but not the additional missions which may appear, could run a similar simulation which incorporates additional missions randomly throughout the iterations. By running a simulation with a slightly perturbed data set the results would be flexible to the uncertainties inherent in mission planning.

4.4 *Training Value Functions and Perturbed Solutions*

In the previous section, the value functions were tested through a series of simulations which looked at how robust the value functions are when faced with varying demands. The results of the previous section illustrate the robustness of the algorithm and the value functions, but they could be considered outside of the realm of possibilities for planning purposes. However, the previous section did highlight the ability of the value functions to incorporate new data on a continual basis and produce acceptable solutions. It is the ability to produce an acceptable solution quickly which will be examined in this section, as it is determined how quickly a perturbed solution can be solved using trained value functions.

During combat mission planning, a mission planner may be tasked with producing a continually updated an aerial refueling solution for inputs which change by the hour. Given the complexities and time required to run a simulation, it could be impossible to continually rerun the refueling model to find a new solution without any shortcuts. This is a common problem in industrial problems when a linear programming approach is required with several hundred thousand or million variables. In an industrial problem, when a linear program is used the fact that a previous solution provides a head start on reaching the optimal solution for a perturbed problem can be exploited. It will be illustrated that this algorithm has a similar structure, such that a perturbed problem can exploit the solutions from a similar problem to quickly converge on a new solution.

This section is not concerned with altering the demands continually throughout the iterations, but rather it focuses on using previously created value functions to quickly find a solution for a perturbed data set. In this manner the perturbations to the inputs can be viewed as perturbing a linear program and using the previous solution as a head start toward reaching optimality. Since the SDS is quickly solved both in iterations required and actual computing time it is not as instructive to use in this simulation and only the LDS will be examined.

To create a new data set, NDS, the LDS was copied so the NDS was twice the size of the LDS. Since the times and requirements of the LDS are already established, it was determined that additional missions in the real world would likely be similar in nature to those of the existing data set. This is due to the requirements facing a mission planner when it is decided that instead of sending four fighters as a bomber escort, six fighter will be sent, or instead of one bomber they will send two bombers and additional fighter escorts.

To test the ability of trained value functions to quickly reach an optimal solution by perturbing the inputs, the first step was to train the value functions through running a 100 iteration simulation on the LDS. After 100 iterations, the inputs were perturbed such that the original LDS missions were included along with a random sample of approximately 20 percent of the LDS missions from the NDS. The simulation was then run for another 50 iterations to determine when a stable solution was reached. As with previous stochastic simulations, a series of simulations were run (five) which were then averaged to get the final results. To further illustrate how the perturbed solutions optimized Figure 63 shows the original optimization of the LDS for 100 iterations along with the perturbed solution which occurs after the 100th iteration.

As shown in Figure 63, by using previously created value functions the aerial refueling model was able to quickly assimilate the new missions. To further illustrate how quickly the model responded, it is illustrative to look at the components of total cost in Figure 65. The receiver's total cost quickly reaches a steady state value as the queuing within the system is brought down to a reasonable level, shown in Figure 64. The tankers take more time to adapt to the new receiver missions, which is due to an overcorrection in response to the increased fueling delays directly after the perturbation. Once the value functions correctly assimilate the new values of having

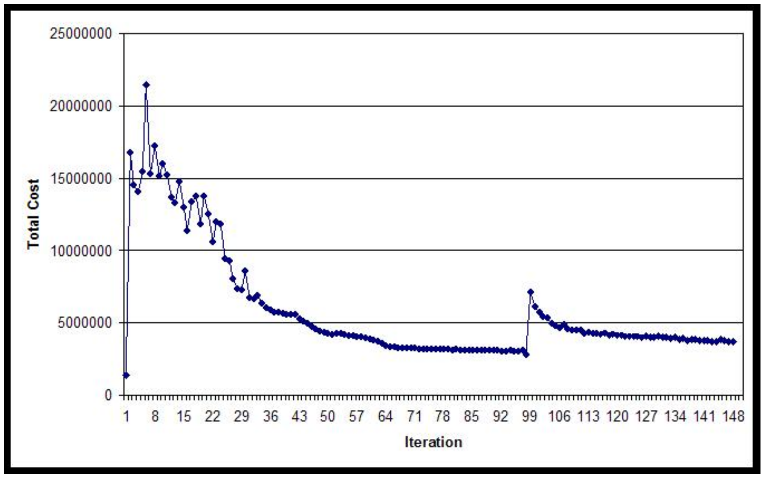

Figure 63: Total Cost for a Data Set (LDS) Perturbed at the 100th Iteration (Adding ≈ 20 Percent More Missions)

additional tankers at a track, the tankers reduce to more natural levels.

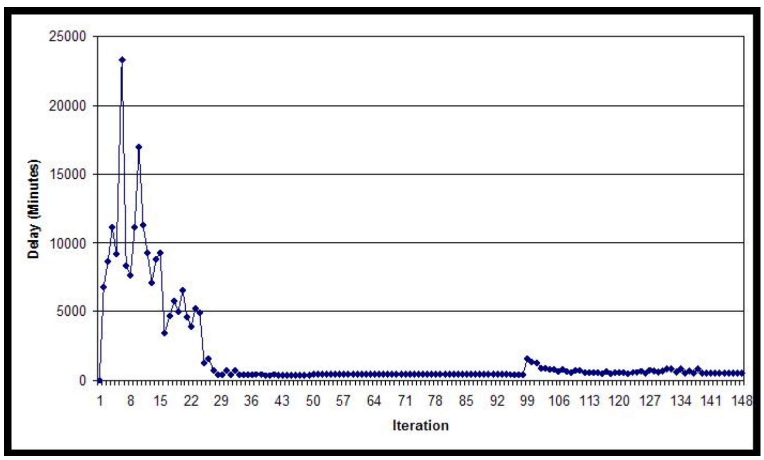

Figure 64: Delay for a Data Set (LDS) Perturbed at the 100th Iteration (Adding ≈ 20 Percent More Missions)

A comparative examination of various outputs from the end of the perturbed simulation (Iteration 150) and the expected values of the outputs (computed as 120% of values at Iteration 100) are shown in Figures 66-69. While the expected values are only approximations as the composition of the perturbed receiver missions entering the system is unknown, it provides a baseline for comparison. Using the expected values as a comparison, the perturbed solution's outputs compare favorably after only

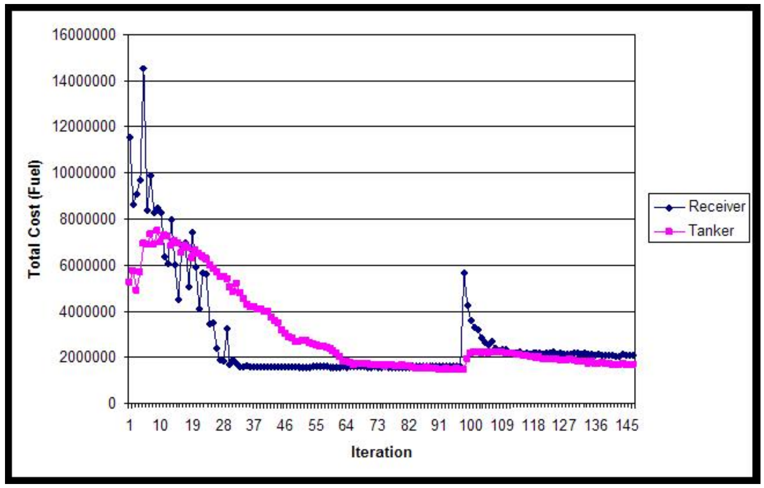

Figure 65: Total Fuel Burned for a Data Set (LDS) Perturbed at the 100th Iteration (Adding ≈ 20 Percent More Missions)

50 iterations. The differences in the delay and tanker fuel cost are lower than their expected values by 7 and 9 percent, while the total cost and receiver cost are higher by 6 and 5 percent, respectively. These values are extremely close and indicate that the model optimized incredibly well with the added mission load. Since the fueling delay is lower than expected but the receiver fuel cost is increased, it indicates that the receiver missions added to the system demanded high fuel loads. Therefore, the cost of refueling those receivers was higher than expected which was reflected in the receiver fuel burn cost and subsequently the total cost of the system.

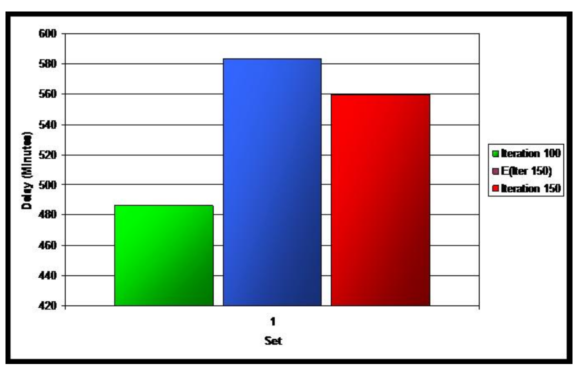

Figure 66: Delay after Perturbation versus Previous Delay and Expected Delay for LDS and Perturbed LDS (Adding ≈ 20 Percent More Missions)

While the previous example of perturbing the data set by 20 percent provided

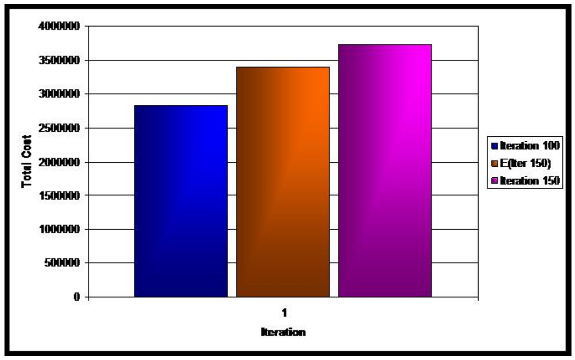

Figure 67: Total Cost after Perturbation versus Previous Cost and Expected Cost for LDS and Perturbed LDS (Adding ≈ 20 Percent More Missions)

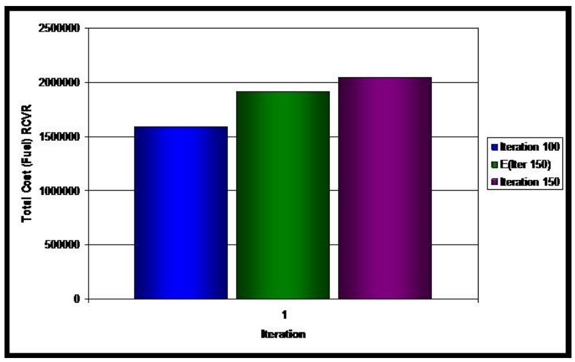

Figure 68: Total Receiver Fuel Cost after Perturbation versus Previous Fuel Cost and Expected Fuel Cost for LDS and Perturbed LDS (Adding ≈ 20 Percent More Missions)

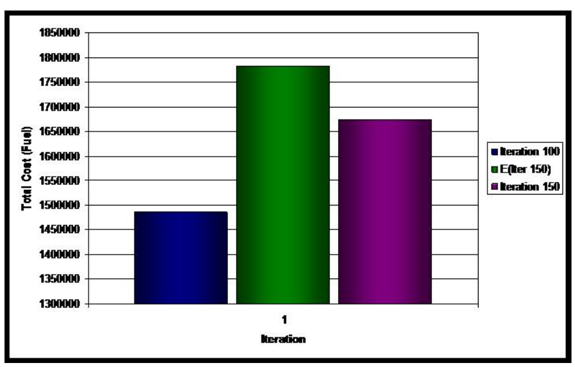

Figure 69: Total Tanker Fuel Cost after Perturbation versus Previous Fuel Cost and Expected Fuel Cost for LDS and Perturbed LDS (Adding ≈ 20 Percent More Missions)

solid results which proved the flexibility and robust nature of the value functions, it was an extreme case. In a more realistic real world example, the perturbations would likely be closer to 5 or 10 percent. To test the ability of the aerial refueling model to assimilate quickly to smaller perturbations, the value functions were trained on the identical data set as before and during the perturbation phase either 5 percent or 10 percent more missions were added to the system.

The results of the smaller perturbations as well as the original perturbation are shown in Figures 70 and 71. For the smaller perturbations the model responds almost immediately in assimilating the missions and reaching an optimal solution. After a brief spike, the value functions are trained to send out the appropriate number of tankers and the total cost settles into a long run value. The smaller perturbations, which are considered to be more realistic, are handled extremely well by the value functions and provide a great deal of value to a mission planner. After doing an initial run a mission planner could store the value functions and respond to any small perturbations by running the perturbed data set with the previously trained value functions. Using previously trained value functions, a mission planner could quickly and accurately assemble all the contingency plans for the days mission or respond on the fly to new mission requirements.

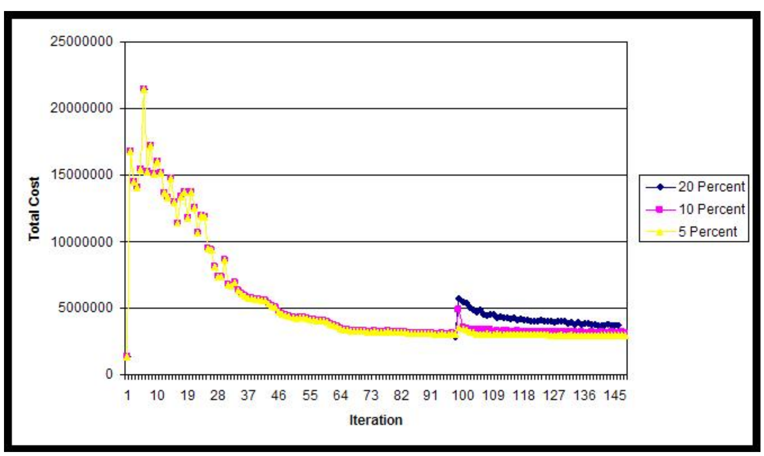

Figure 70: Testing Different Levels of Perturbation and Their Rates of Convergence (Total Cost) after the Perturbations

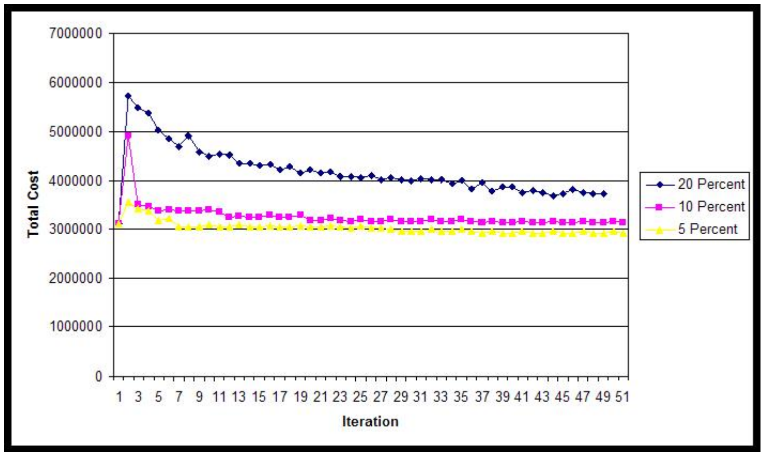

Figure 71: Testing Different Levels of Perturbation and Their Rates of Convergence (Total Cost) after the Perturbations

The capabilities of the aerial refueling model to assimilate stochastic data are of great use to Air Force mission planners. The ability to quickly respond to the frictions of warfare and produce usable results is a major strength of the model. The cornerstone to the flexibility of the model are the value functions which in the stochastic sections of this thesis have been proven to be very robust. The value functions have been shown to accommodate uncertainties of fuel loads, refueling times, and most impressively differing receiver mission inputs. The ability of the value functions to adapt to different stochastic inputs is a great strength of the model which cannot be replicated in a myopic simulation model and could provide the Air Force with an increased ability to plan combat missions.

5 *Conclusions*

The ability of the aerial refueling model to accurately model the realities of in-flight refueling are a leaps and bounds improvement over the current system. The model is relatively insensitive to inputs in the system such as tankers and provides incredibly robust solutions. The solution quality produced by the aerial refueling model is both efficient as well as flexible, which is a hallmark of solutions produced through approximate dynamic programming.

Continuing refinement and expansion of the aerial refueling model could provide a boon for the capabilities of the modern US Air Force fleet. Through the use of the aerial refueling model the existing capabilities of the refueling fleet can be expanded and support combat operations for the foreseeable future.

References

[1] AVRO VULCAN. Thunder-and-lightnings.co.uk British Aviation Heritage, Midland Air Museum and the North East Aircraft Museum. June, 2006. 11, 13

[2] Bolkcom, Christopher. Air Force Aerial Refueling Methods: Flying Boom versuses Hose-and-Drogue. CRS Reort for Congress (RL32910). June 5, 2006. 8, 17

[3] Curtin, Neal. Military aircraft: Information on air force refueling tankers. (GAO-03-938T), June 2003. 17

[4] Defense Science Board Task Force Report. AERIAL REFUELING REQUIREMENTS. May 2004. 17

[5] Day, Dwayne. *Aerial Refueling*. U.S. Centennial of Flight Commision. $centennialofflight.gov/essay/Evolution_of_Technology/refueling?Tech22.htm$. 20003. 5

[6] DOD Needs to Determine It Aerial Refueling Aircraft Requirements. United States General Accounting Office. June, 2004.

[7] Gehma, Harold Jr., Admiral, USN (Ret) Progress Report on Joint Experimentation. Joint Forces Quarterly. Summer, 2000. 1

[8] Grant, Rebecca. The Afghan Air War. Aerospace Education Foundation. September, 2002. 16

[9] Harvey, David W. *Seventy Years of Strategic Air Refueling, 1918-1988: a Chronology.* USAF, Offutt AFB, Nebraska, May 1990. 5, 7, 9, 11

[10] Jumper, John, General, USAF. Global Strike Task Foce- *A Transforming Concept, Forged by Experience.* Aerospace Power Journal. Spring, 2001. 14

[11] Killingsworth, Paul S. *Multipoint Aerial Refueling: A Review and Assesment.* RAND Corp., Santa Monica, 1996. 15

[12] Lepore, Brian. Military aircraft: Dod needs to determine its aerial refueling requirements. (GAO-04-359), November 1996. 14

[13] Lexington Institute. Modernizing The Aerial Refueling Fleet. Arlington, Virginia. September 2006. 7

[14] Ochinko, Walter. Aerial refueling initiative: Cross service analysis needed to determine best approach. (GAO/NSIAD-94-68), November 1996. 14

[15] Ochinko, Walter. Operation desert storm: An assessment of aerial refueling operations efficiency. (GAO/NSIAD-94-68), November 1996. 15

[16] Pope, Charles. Boeing tanker deal faces inquiry. September 2003.

[17] Powell, Warren. Approximate Dynamic Programming. Unpublished as of May 2007. 20, 24, 25, 26, 110

[18] Powell, W. and George, A. Adaptive stepsiezes for recursive estimation with applications in approximate dynamic programming. Mach Learn(2006) 65:167-198 51

[19] Ritter, Douglas. *Aerial Refueling* Code One: An Airpower Projection Magazine. January 1993. 5

[20] Smith, Richard. *Seventy-Five Years of Inflight Refueling: Highlights, 1923-1928.* USAF, Washington, DC., 1998. 6, 7, 9, 10, 13

[21] Spiegel, Peter. Aerial Tanker Goes to Top of List; The Air Force makes the acquistion of refuelers, possibly from Boeing, its No. 1 priority. Los Angeles Times. October 13, 2006. 17

[22] White, Rowland. Vulcan 607. London. Bantam, 2006. vii, 12